

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
7 October 2004 (07.10.2004)

PCT

(10) International Publication Number  
**WO 2004/085656 A2**

(51) International Patent Classification<sup>7</sup>: C12N 15/82,  
15/53, 15/54, 15/62, A01H 5/10, 5/00, C07C 403/00

(74) Agents: SOUTHERN, David, William et al.; Intellectual  
Property Department, Syngenta Limited, P.O. Box 3538,  
Jealott's Hill International Research Centre, Bracknell,  
Berkshire RG42 6YA (GB).

(21) International Application Number:  
PCT/GB2004/001241

(22) International Filing Date: 22 March 2004 (22.03.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/457,053 24 March 2003 (24.03.2003) US

(71) Applicant (for all designated States except US): SYN-  
GENTA LIMITED [GB/GB]; European Regional Centre,  
Priestley Road, Surrey Research Park, Guildford, Surrey  
GU2 7YH (GB).

(72) Inventors; and

(75) Inventors/Applicants (for US only): DRAKE, Caro-  
line, Rachel [GB/GB]; Syngenta Limited, Jealott's Hill  
International Research Centre, Bracknell, Berkshire RG42  
6EY (GB). PAINE, Jacqueline, Ann, Mary [GB/GB];  
Syngenta Limited, Jealott's Hill International Research  
Centre, Bracknell, Berkshire RG42 6EY (GB). SHIPTON,  
Catherine, Ann [GB/GB]; Syngenta Limited, Jealott's  
Hill International Research Centre, Bracknell, Berkshire RG42  
6EY (GB).

(81) Designated States (unless otherwise indicated, for every  
kind of national protection available): AE, AG, AL, AM,  
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,  
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI,  
GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,  
KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD,  
MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG,  
PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM,  
TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM,  
ZW.

(84) Designated States (unless otherwise indicated, for every  
kind of regional protection available): ARIPO (BW, GH,  
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), Euro-  
pean (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR,  
GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK,  
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,  
ML, MR, NE, SN, TD, TG).

**Published:**

— without international search report and to be republished  
upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.

(54) Title: ENHANCED ACCUMULATION OF CAROTENOIDS IN PLANTS

(57) Abstract: The present invention relates to polynucleotides and their use in methods of increasing the carotenoid content of seeds. In particular the invention provides a polynucleotide comprising: (a) a region which comprises as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a nucleotide sequence derived from a bacterium which sequence encodes a carotene desaturase; and (iii) a transcription termination region; and (b) a further region which comprises as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a nucleotide sequence encoding a phytoene synthase which sequence is derived from maize (*Zea* sp.) or rice (*Oryza* sp.); and (iii) a transcription termination region. The disclosed polynucleotides are particularly suitable for use in production of rice seed which comprise high amounts of coloured carotenoids.



WO 2004/085656 A2

$$z/PRf_s^{-1}$$

## ENHANCED ACCUMULATION OF CAROTENOIDS IN PLANTS

The present invention relates *inter alia*, to recombinant DNA technology. More specifically the invention relates to the provision of improved polynucleotides which provide for an enhanced accumulation of carotenoids in plants and in particular in the seeds of said plants. The invention also provides plant material, plants and seeds which comprise the polynucleotides, in particular rice plant material, rice plants and rice seeds.

Carotenoids are 40-carbon ( $C_{40}$ ) isoprenoids formed by condensation of eight isoprene units derived from the biosynthetic precursor isopentenyl diphosphate. By nomenclature, carotenoids basically fall into two classes, namely, carotenes and xanthophylls. Their essential function in plants is to protect against photo-oxidative damage in the photosynthetic apparatus of plastids. In addition to this, carotenoids participate in light harvesting during photosynthesis and represent integral components of photosynthetic reaction centres. Carotenoids are the direct precursors of the phytohormone abscisic acid. A part of the carotenoid biosynthesis pathway is shown in Figure 1.

Carotenoids with provitamin A activity are essential components of the human diet. Additionally, there is compelling evidence suggesting that a diet rich in carotenoids can prevent a number of serious medical conditions from developing, including certain cancers (especially lung and prostate), macular degeneration, cataract and cardiovascular disease. Carotenoids have also been reported to have immunomodulatory effects, such as the reduction in UV-induced immunosuppression. Carotenoids are able to act as efficient quenchers of harmful reactive oxygen species such as singlet oxygen and thereby have antioxidant properties.

25           With the population of the world increasing, there remains a need for the  
production of foods which are high in nutrition, healthy, tasty and visually appealing.

The general insertion of genes involved in the carotenoid biosynthesis pathway into plants is disclosed in WO00/53768. US6,429,356 describes methods for the production of plants and seeds having altered fatty acid, tocopherol and carotenoid compositions via insertion of a crtB gene. The present invention provides, *inter alia*, improved polynucleotides which when inserted into plant material provide for a surprisingly high accumulation of carotenoids in at least the seeds of plants derived from said material. More specifically, the present invention provides particular combinations

- 2 -

of nucleotide sequences which, when expressed in plant material provide for a surprisingly high accumulation of carotenoids in at least the seeds of plants derived from said material.

According to the present invention there is provided a polynucleotide comprising:

- 5 (a) a region which comprises as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a nucleotide sequence derived from a bacterium which sequence encodes a carotene desaturase; and (iii) a transcription termination region; and (b) a further region which comprises as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a  
10 nucleotide sequence encoding a phytoene synthase which sequence is derived from maize (*Zea* sp.) or rice (*Orzya* sp.); and (iii) a transcription termination region.

The invention further provides a polynucleotide comprising: (a) a region which comprises as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a nucleotide sequence derived from a bacterium which  
15 sequence encodes a carotene desaturase; and (iii) a transcription termination region; and (b) a further region which comprises as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a nucleotide sequence encoding a phytoene synthase which sequence is derived from tomato (*Lycopersicon* sp.) or pepper (*Capsicum* sp.) or a bacterium; and (iii) a transcription termination region.

20 In a further embodiment said carotene desaturase is derivable from *Streptomyces*; *Staphylococcus*; *Synechocystis*; *Rhodobacter*; *Paracoccus*; *Erwinia*; and *Xanthophyllomyces*. In a further embodiment said carotene desaturase is a phytoene desaturase. In a further embodiment when said phytoene synthase is derivable from a bacterium, it is derivable from *Streptomyces*; *Staphylococcus*; *Synechocystis*;  
25 *Rhodobacter*; *Paracoccus*; *Erwinia*; and *Xanthophyllomyces*. In a further embodiment of the invention said phytoene synthase is obtainable from *Zea mays* or *Orzya sativa*. In a still further embodiment of the invention said phytoene synthase is obtainable from *Lycopersicon esculentum* or *Capsicum annuum*.

In a further embodiment said phytoene synthase is obtained from *Zea mays* or  
30 *Orzya sativa*. In a still further embodiment said phytoene synthase comprises, is comprised by or consists of the sequence selected from the group depicted as SEQ ID NOS: 10; 11; 12; and 13. In a further embodiment said phytoene synthase comprises, is comprised by or consists of a sequence which encodes the protein depicted as SEQ ID

NO 14. Alternative phytoene synthase encoding sequences may also be available from databases known to the person skilled in the art. For example, the sequences depicted in the EMBL Database under Accession Numbers – AY024351, AK073290, AK108154 and AY078162.

5 In a still further embodiment said phytoene synthase is derived, obtainable or obtained from *Lycopersicon esculentum* or *Capsicum annuum*. In a still further embodiment said phytoene synthase comprises, is comprised by or consists of SEQ ID NO: 15 or SEQ ID NO: 16.

10 In a still further embodiment said phytoene synthase is derived, obtainable or obtained from the CrtB gene from bacteria. In a particular embodiment said crtB comprises, is comprised by or consists of the crtB sequence depicted in Shewmaker et al. (1999) Plant J. 20:401-12. In a further embodiment of the invention the sequence which encodes a carotene desaturase and/or the sequence encoding a phytoene synthase from a bacteria is derived from *Erwinia sp.*, more specifically, *Erwinia uredovora*. In a still  
15 further embodiment the carotene desaturase is derived from the CrtI gene from *Erwinia uredovora*. In a still further embodiment the carotene desaturase is the CrtI gene from *Erwinia uredovora*. In a still further embodiment the carotene desaturase comprises, is comprised by or consists of the sequence depicted as SEQ ID NO: 18 or SEQ ID NO: 19.

The present invention further provides a polynucleotide as described above  
20 wherein said promoter is selected from the Glutelin 1 promoter and the Prolamin promoter and said transcription termination region is selected from the Nos; CaMV 35S and PotP1-II transcription termination regions. Said Glutelin 1 and Prolamin promoters may be isolated from rice. In a particular embodiment said promoter comprises the sequence depicted as SEQ ID NO: 20 or 21. Further promoters include the promoter  
25 derived from the napin gene from *Brassica napus* and other promoters which are derived from genes normally expressed in the endosperm of the seed. Further transcription termination regions include the terminator region of a gene of alpha-tubulin (EP-A 652,286). It is equally possible to use, in association with the promoter regulation sequence, other regulation sequences which are situated between the promoter and the  
30 sequence encoding the protein, such as transcriptional or translational enhancers, for example, tobacco etch virus (TEV) translation activator described in International Patent application, PCT publication number WO87/07644.

The present invention still further provides a polynucleotide as described above wherein the sequence which encodes carotene desaturase and the sequence which encodes phytoene synthase further comprises a plastid targeting sequence. In a particular embodiment the plastid targeting sequence is derived from the ribulose biphosphate carboxylase small-subunit (RUBISCO Ssu) from *Pisum sativum*. In a further embodiment said RUBISCO Ssu plastid targeting sequence is located 5' to the translational start point of said carotene desaturase gene derived from CrtI. In a still further embodiment said RUBISCO Ssu plastid targeting sequence is located 5' to the translational start point of said phytoene synthase gene derived from CrtB. In a further embodiment said plastid targeting sequence is heterologous with respect to said phytoene synthase and/or said carotene desaturase. In a still further embodiment said plastid targeting sequence is autologous with respect to said phytoene synthase and/or said carotene desaturase. By "heterologous" is meant from a different source, and correspondingly "autologous" means from the same source - but at a gene rather than organism or tissue level. In a still further embodiment the plastid targeting sequence associated with the carotene desaturase is heterologous therewith and the plastid targeting sequence associated with the phytoene synthase is autologous therewith. In a still further embodiment the plastid targeting sequence provides for the accumulation of carotenoids in the amyloplast of the seed.

The present invention still further provides a polynucleotide as described above wherein the sequence which encodes said carotene desaturase and/or the sequence which encodes said phytoene synthase further comprises an intron. In a particular embodiment said intron region is located between the promoter region and the region encoding the carotene desaturase/phytoene synthase. In a further embodiment said intron region is located between the promoter region and the plastid targeting sequence. In a still further embodiment said intron region is located upstream of said carotene desaturase and/or said phytoene synthase encoding sequence(s). In a still further embodiment said intron is derived from the intron of the first gene from the catalase gene of the castor bean plant. In a still further embodiment said intron is from the first intron of the catalase gene from the castor bean plant. In a still further embodiment said intron comprises the sequence depicted as SEQ ID NO: 22. In a still further embodiment the intron is the intron of the maize polyubiquitin gene.

The present invention further provides a polynucleotide as described above wherein said sequence encoding carotene desaturase is located 5' to said sequence encoding phytoene synthase.

5 The present invention further provides a polynucleotide as described above wherein said sequence encoding phytoene synthase is located 5' to said sequence encoding carotene desaturase.

The present invention still further provides a polynucleotide as described above which comprises the sequence selected from the group depicted as SEQ ID NOS: 1; 2; 3; 4; 5; and 6. In a particular embodiment of the invention the polynucleotide consists of a  
10 sequence selected from the group depicted as SEQ ID NO: 1; 2; 3; 4; 5 and 6. In a further embodiment of the invention the polynucleotide comprises or consists of SEQ ID NO: 1. In a still further embodiment of the invention the polynucleotide comprises or consists of SEQ ID NO: 2. In a still further embodiment of the invention the polynucleotide comprises or consists of SEQ ID NO: 3. In a still further embodiment of  
15 the invention the polynucleotide comprises or consists of SEQ ID NO: 4. In a still further embodiment of the invention the polynucleotide comprises or consists of SEQ ID NO: 6. The present invention still further provides a polynucleotide as described above which comprises or consists of a sequence selected from the group depicted as SEQ ID NOS: 7; 8; and 9.

20 The present invention still further provides a polynucleotide sequence which is the complement of one which hybridises to a polynucleotide as described in the preceding paragraph at a temperature of about 65°C in a solution containing 6 x SSC, 0.01% SDS and 0.25% skimmed milk powder, followed by rinsing at the same temperature in a solution containing 0.2 x SSC and 0.1% SDS wherein said  
25 polynucleotide sequence still comprises a region encoding a carotene desaturase and a further region encoding a phytoene synthase and when said polynucleotide sequence is inserted into plant material the seed of a plant regenerated from said material produce an increased amount of carotenoids when compared to a control like-seed. The skilled person may alternatively select the following hybridisation conditions, viz., hybridisation  
30 at a temperature of between 60°C and 65°C in 0.3 strength citrate buffered saline containing 0.1% SDS followed by rinsing at the same temperature with 0.3 strength citrate buffered saline containing 0.1% SDS followed by confirmation that when the polynucleotide sequence so identified is inserted into plant material the seed of a plant

regenerated from said material produce an increased amount of carotenoids when compared to a control like-seed. The person skilled in the art may also select further hybridisation conditions that are equally understood to be "high stringency" conditions. In a particular embodiment of the present invention when said polynucleotide sequence is inserted into plant material, the seed of a plant regenerated from said material produces at least a sixty fold increase in carotenoids when compared to a control like-seed. In a further embodiment of the invention when said polynucleotide sequence is inserted into plant material, the seed of a plant regenerated from said material produces at least a one hundred fold increase in carotenoids when compared to a control like-seed. In a still further embodiment of the invention when said polynucleotide sequence is inserted into plant material, the seed of a plant regenerated from said material produces at least a two hundred and fifty fold increase in carotenoids when compared to a control like-seed. In a still further embodiment of the invention when said polynucleotide sequence is inserted into plant material, the seed of a plant regenerated from said material produces at least a three hundred and fifty fold increase in carotenoids when compared to a control like-seed. In a still further embodiment of the invention when said polynucleotide sequence is inserted into plant material, the seed of a plant regenerated from said material produces at least a four hundred fold increase in carotenoids when compared to a control like-seed. In a still further embodiment of the invention when said polynucleotide sequence is inserted into plant material, the seed of a plant regenerated from said material produces at least a five hundred fold increase in carotenoids when compared to a control like-seed.

The term control like-seed relates to seeds which are substantially similar to those according to the invention but which control like-seed does not contain the polynucleotides or polynucleotide sequences according to the invention. Typically, a

control like-seed will comprise a seed of the same or similar plant species which control like-seed has not been transformed. The increased carotenoid content of the seeds comprising the polynucleotides or polynucleotide sequences according to the invention may also be demonstrated via comparison of said seeds with seeds that comprise the TDNA depicted in Plasmid A of Figure 4 of WO00/53768, wherein the phytoene synthase (psy) is from daffodil (*Narcissus pseudonarcissus*). Typically, such a comparison would be made when the seed to be compared are of the same or a substantially similar plant species. In a particular embodiment the seed comprising the polynucleotides or polynucleotide sequences according to the invention contain at least three times the amount of carotenoids when compared to a seed that comprise the TDNA depicted in Plasmid A of Figure 4 of WO00/53768, wherein the phytoene synthase (psy) is from daffodil (*Narcissus pseudonarcissus*). In a still further embodiment the seed comprising the polynucleotides or polynucleotide sequences according to the invention contains at least four times, or at least five times, or at least six times, or at least seven times, or at least eight times or at least nine times or at least ten times or at least fifteen times or at least twenty times or at least thirty times, or at least forty times or at least fifty times the amount of carotenoids when compared to a seed that comprise the TDNA depicted in Plasmid A of Figure 4 of WO00/53768, wherein the phytoene synthase (psy) is from daffodil (*Narcissus pseudonarcissus*).

The present invention still further provides a polynucleotide sequence as described above wherein when said polynucleotide sequence is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least 3µg/g of endosperm of said seed. In a further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least 4µg/g of endosperm of said seed. In a further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least 5µg/g of endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least 6µg/g of endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a



level of at least  $7\mu\text{g/g}$  of endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least  $8\mu\text{g/g}$  of endosperm of said seed. In a still further embodiment when said polynucleotide

5 sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least  $9\mu\text{g/g}$  of endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least  $10\mu\text{g/g}$  of

10 endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least  $11\mu\text{g/g}$  of endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant

15 regenerated from said material produces carotenoids at a level of at least  $12\mu\text{g/g}$  of endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least  $13\mu\text{g/g}$  of

20 endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least  $14\mu\text{g/g}$  of endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant

25 regenerated from said material produces carotenoids at a level of at least  $15\mu\text{g/g}$  of endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least  $20\mu\text{g/g}$  of

30 endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least  $25\mu\text{g/g}$  of endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces

carotenoids at a level of at least 30 $\mu$ g/g of endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least 35 $\mu$ g/g of endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least 40 $\mu$ g/g of endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least 45 $\mu$ g/g of endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least 50 $\mu$ g/g of endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least 55 $\mu$ g/g of endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least 60 $\mu$ g/g of endosperm of said seed. In a still further embodiment when said polynucleotide sequence as described above is inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at a level of at least 65 $\mu$ g/g of endosperm of said seed. In a particular embodiment the amount of carotenoids is calculated as  $\mu$ g/g of dry weight of endosperm of said seed.

The present invention still further provides a polynucleotide sequence which is the complement of one which hybridises to a polynucleotide which consists of a sequence selected from the group depicted as SEQ ID NOS: 1; 2; 3; 4; 5; and 6 at a temperature of about 65°C in a solution containing 6 x SSC, 0.01% SDS and 0.25% skimmed milk powder, followed by rinsing at the same temperature in a solution containing 0.2 x SSC and 0.1% SDS wherein said polynucleotide sequence still comprises a region encoding a carotene desaturase and a further region encoding a phytoene synthase and when said polynucleotide sequence is inserted into plant material the seed of a plant regenerated from said material produce carotenoids amounting to at least 80% of the carotenoid content of a seed which comprises a polynucleotide selected

SEQ ID NOS: 1; 2; 3; 4; 5 and 6. In a further embodiment, a polynucleotide sequence is inserted into plant material, the seed of a plant regenerated from said material produce carotenoids amounting to at least 85% of the carotenoid content of a seed which comprises a polynucleotide selected from the group depicted as SEQ ID NOS: 1; 2; 3; 4; 5 and 6. In a still further embodiment when said polynucleotide sequence is inserted into plant material, the seed of a plant regenerated from said material produce carotenoids amounting to at least 90% of the carotenoid content of a seed which comprises a polynucleotide selected from the group depicted as SEQ ID NOS: 1; 2; 3; 4; 5; and 6. In a still further embodiment when said polynucleotide sequence is inserted into plant material, the seed of a plant regenerated from said material produce carotenoids amounting to at least 95% of the carotenoid content of a seed which comprises a polynucleotide selected from the group depicted as SEQ ID NOS: 1; 2; 3; 4; 5 and 6. In a still further embodiment when said polynucleotide sequence is inserted into plant material, the seed of a plant regenerated from said material produce carotenoids amounting to at least 100% of the carotenoid content of a seed which comprises a polynucleotide selected from the group depicted as SEQ ID NOS: 1; 2; 3; 4; 5 and 6. In a particular embodiment the polynucleotide sequence provides for a percentage of the carotenoid content of seed as described above wherein the seed with which the comparison is made comprises the polynucleotide depicted as SEQ ID NO: 1. In a particular embodiment the polynucleotide sequence provides for a percentage of the carotenoid content of seed as described above wherein the seed with which the comparison is made comprises the polynucleotide depicted as SEQ ID NO: 2. In a particular embodiment the polynucleotide sequence provides for a percentage of the carotenoid content of seed as described above wherein the seed with which the comparison is made comprises the polynucleotide depicted as SEQ ID NO: 3. In a particular embodiment the polynucleotide sequence provides for a percentage of the carotenoid content of seed as described above wherein the seed with which the comparison is made comprises the polynucleotide depicted as SEQ ID NO: 4. In a particular embodiment the polynucleotide sequence provides for a percentage of the carotenoid content of seed as described above wherein the seed with which the comparison is made comprises the polynucleotide depicted as SEQ ID NO: 6.

It is preferred that when the carotenoid content of the seed comprising said polynucleotide sequence is compared with the seed comprising the polynucleotide

selected from the group depicted as SEQ ID NOS: 1; 2; 3; 4; 5 and 6, the seed are from plants of substantially the same species. It is further preferred that when the carotenoid content of the seed comprising said polynucleotide sequence is compared with the seed comprising the polynucleotide selected from the group depicted as SEQ ID NOS: 1; 2; 3; 4; 5 and 6, the seed are from plants which are substantially genetically identical save for the presence of said polynucleotide or said polynucleotide sequence. It is further preferred that when the carotenoid content of the seed comprising said polynucleotide sequence is compared with the seed comprising the polynucleotide selected from the group depicted as SEQ ID NOS: 1; 2; 3; 4; 5 and 6, the seed are from plants which are grown subject to the same environmental growing conditions.

The present invention still further provides a polynucleotide sequence which is the complement of one which hybridises to a polynucleotide which consists of a sequence selected from the group depicted as SEQ ID NOS: 7; 8; and 9 at a temperature of about 65°C in a solution containing 6 x SSC, 0.01% SDS and 0.25% skimmed milk powder, followed by rinsing at the same temperature in a solution containing 0.2 x SSC and 0.1% SDS wherein said polynucleotide sequence still comprises a region encoding a carotene desaturase and a further region encoding a phytoene synthase and when said polynucleotide sequence is inserted into plant material the seed of a plant regenerated from said material produce carotenoids amounting to at least 80% of the carotenoid content of a seed which comprises a polynucleotide selected from the group depicted as SEQ ID NOS: 7; 8 and 9. It is preferred that when the carotenoid content of the seed comprising said polynucleotide sequence is compared with the seed comprising the polynucleotide selected from the group depicted as SEQ ID NOS: 7; 8 and 9, the seed are from plants of substantially the same species. It is further preferred that when the carotenoid content of the seed comprising said polynucleotide sequence is compared with the seed comprising the polynucleotide selected from the group depicted as SEQ ID NOS: 7; 8 and 9, the seed are from plants which are substantially genetically identical save for the presence of said polynucleotide or said polynucleotide sequence.

In a particular embodiment the polynucleotide sequences according to the invention which are identified based on their hybridisation (under the conditions provided) to the sequences described in the Sequence Listing, encode the same proteins as those provided by the sequences in the Sequence Listing. In a further embodiment, said polynucleotide sequences may encode proteins which have the same or a similar

function as the proteins encoded by the sequences in the Sequence Listing. In a still further embodiment, the proteins encoded by the polynucleotide sequence according to the invention comprise amino acid substitutions and/or deletions when compared to the proteins encoded by the sequences in the Sequence Listing. In a still further  
5 embodiment, said amino acid substitutions are "conservative" substitutions. A "conservative" substitution is understood to mean that the amino acid is replaced with an amino acid with broadly similar chemical properties. In particular "conservative" substitutions may be made between amino acids within the following groups: (i) Alanine and Glycine; (ii) Threonine and Serine; (ii) Glutamic acid and Aspartic acid; (iii)  
10 Arginine and Lysine; (iv) Asparagine and Glutamine; (v) Isoleucine and Leucine; (vi) Valine and Methionine; and (vii) Phenylalanine and Tryptophan.

The present invention still further provides a polynucleotide or a polynucleotide sequence as described above which when inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at levels which are higher than those  
15 present in native like-seeds. The present invention still further provides a polynucleotide or a polynucleotide sequence as described above which when inserted into plant material, the seed of a plant regenerated from said material produces carotenoids at levels which are higher than those present in untransformed like-seeds.

In a particular embodiment, the carotenoids which are increased are selected from  
20 the group consisting of: lycopene; alpha-carotene; lutein; beta-carotene; zeaxanthin; beta-cryptoxanthin; antheraxanthin; violaxanthin; and neoxanthin or a combination thereof. In a further embodiment, the carotenoids which are increased are selected from the group consisting of: lycopene; alpha-carotene; lutein; beta-carotene; zeaxanthin; beta-cryptoxanthin; or a combination thereof. In a still further embodiment, the carotenoids  
25 which are increased are selected from the group consisting of: alpha-carotene; lutein; beta-carotene; zeaxanthin; beta-cryptoxanthin; or a combination thereof. In a still further embodiment, the carotenoids which are increased include at least phytoene and beta-carotene. In a still further embodiment, the carotenoids which are increased include at least beta-carotene. In a still further embodiment, the carotenoid which is increased is  
30 beta-carotene. In a still further embodiment the carotenoids which are increased are coloured carotenoids.

The present invention still further provides a polynucleotide or a polynucleotide sequence as described above wherein said seed is a rice seed. In a particular embodiment

of the invention, before the seeds are analysed for their carotenoid content, the seeds are prepared prior to the analysis. Such preparation may include, for example, with respect to rice seed, "dehusking" and "polishing". Furthermore, such preparation may involve the removal of those plant parts associated with the seed which plant parts are not  
5 normally intended for human consumption.

The amount of carotenoids in the seeds can be determined using techniques which are well known and available to the person skilled in the art. Such techniques include but are not necessarily limited to High Performance Liquid Chromatography (HPLC) analysis and spectrophotometry.

10 The present invention still further provides a polynucleotide or polynucleotide sequence as described above which further comprises a region which encodes a selectable marker. In a particular embodiment said selectable marker comprises a mannose-6-phosphate isomerase gene. In a further particular embodiment the selectable marker used is the mannose-6-phosphate isomerase gene according to the Positech<sup>TM</sup>  
15 selection system. In a specific embodiment said selectable marker is the one as described in European Patent/Application publication Number EP 0 896 063 and EP 0 601 092. Alternatively, the selectable marker used may, in particular, confer resistance to kanamycin, hygromycin or gentamycin. Further suitable selectable markers include genes that confer resistance to herbicides such as glyphosate-based herbicides (e.g.  
20 EPSPS genes such as in USP 5510471 or WO 00/66748) or resistance to toxins such as eutypine. Other forms of selection are also available such as hormone based selection systems such as the Multi Auto Transformation (MAT) system of Hiroyasu Ebinuma *et al.* 1997. PNAS Vol. 94 pp2117-2121; visual selection systems which use fluorescent proteins,  $\beta$  glucuronidase and any other selection system such as xylose isomerase and 2-  
25 deoxyglucose (2-DOG).

The present invention still further provides a polynucleotide or a polynucleotide sequence according to the invention which is codon optimised for expression in a particular plant species. In a particular embodiment the polynucleotide or polynucleotide sequence is codon optimised for expression in rice (*Oryza sp.*) or maize (*Zea sp.*). Such  
30 codon optimisation is well known to the person skilled in the art and the table below provides an example of the plant-preferred codons for rice and maize.

- 14 -

Amino Acid	Rice preference	Maize preference
Alanine	GCC	GCC
Arginine	CGC	AGG
Asparagine	AAC	ACC
Aspartic Acid	GAC	GAC
Cysteine	TGC	TGC
Glutamine	CAG	CAG
Glutamic Acid	GAG	GAG
Glycine	GGC	GGC
Histidine	CAC	CAC
Isoleucine	ATC	ATC
Leucine	CTC	CTG
Lysine	AAG	AAG
Methionine	ATG	ATG
Phenylalanine	TTC	TTC
Proline	CCG	CCG
Serine	TCC	AGC
Threonine	ACC	ACC
Tryptophan	TGG	TGG
Tyrosine	TAC	TAC
Valine	GTG	GTG

The present invention further provides a vector comprising a polynucleotide or a polynucleotide sequence as described above. In a particular embodiment of the invention said vector comprises a polynucleotide selected from the group depicted as SEQ ID NO: 1; 2; 3; 4; 5 and 6. In a particular embodiment of the invention said vector comprises a polynucleotide selected from the group depicted as SEQ ID NOS: 7; 8 and 9. In a particular embodiment said vector allows for replication of said polynucleotide or polynucleotide sequence in a bacterium.

The present invention still further provides a vector as described above which is a plant expression vector.

In a particular embodiment of the invention the sequence around the translational start position(s) of the phytoene synthase and/or said carotene desaturase encoding sequences as described above may be modified such that it is "Kozak" preferred. What is meant by this is well known to the skilled artisan. Examples of Kozak consensus sequences which are well known to the person skilled in the art include cagcc(atg) or agcc(atg). The phytoene synthase and/or said carotene desaturase encoding sequences as described above may also further comprise a sequence which provides for retention in a particular intracellular organelle.

In a further aspect of the present invention there is provided a method for increasing the carotenoid content of seeds comprising inserting into plant material a polynucleotide or a polynucleotide sequence or a vector as described above; and regenerating a seed-containing plant from said material and identifying the seeds which contain carotenoids at levels greater than those of a control like-seed.

The present invention still further provides a method for increasing the carotenoid content of seeds comprising inserting into plant material a polynucleotide comprising a sequence selected from the group depicted as SEQ ID NO: 1; 2; 3; 4; 5 and 6 and regenerating a seed containing plant from said material and identifying the seeds which contain carotenoids at levels greater than those of control like-seeds. In a particular embodiment of the invention, the seeds obtained by said method contain at least a sixty fold increase in carotenoids when compared to a control like-seed. In a further embodiment the seeds obtained by said method contain at least a one hundred fold increase in carotenoids when compared to control like-seeds. In a further embodiment the seeds obtained by said method contain at least a one hundred and fifty fold increase in carotenoids when compared to control like-seeds. In a further embodiment the seeds obtained by said method contain at least a two hundred fold increase in carotenoids when compared to control like-seeds. In a further embodiment the seeds obtained by said method contain at least a two hundred and fifty fold increase in carotenoids when compared to control like-seeds. In a further embodiment the seeds obtained by said method contain at least a three hundred fold increase in carotenoids when compared to control like-seeds. In a further embodiment the seeds obtained by said method contain at least a three hundred and fifty fold increase in carotenoids when compared to control like-seeds. In a further embodiment the seeds obtained by said method contain at least a four hundred fold increase in carotenoids when compared to control like-seeds. In a



further embodiment the seeds obtained by said method contain at least a five hundred fold increase in carotenoids when compared to control like-seeds.

The present invention still further provides a method as described above wherein said seed contains carotenoids at a level of at least 3 $\mu$ g/g of endosperm of said seed. In a particular embodiment said seed contains carotenoids at a level of at least 4 $\mu$ g/g of endosperm of said seed. In a particular embodiment said seed contains carotenoids at a level of at least 5 $\mu$ g/g of endosperm of said seed. In a particular embodiment said seed contains carotenoids at a level of at least 6 $\mu$ g/g of endosperm of said seed. In a particular embodiment said seed contains carotenoids at a level of at least 7 $\mu$ g/g of endosperm of said seed. In a particular embodiment said seed contains carotenoids at a level of at least 8 $\mu$ g/g of endosperm of said seed. In a particular embodiment said seed contains carotenoids at a level of at least 9 $\mu$ g/g of endosperm of said seed. In a particular embodiment said seed contains carotenoids at a level of at least 10 $\mu$ g/g of endosperm of said seed. In a particular embodiment said seed contains carotenoids at a level of at least 15 $\mu$ g/g of endosperm of said seed. In a particular embodiment said seed contains carotenoids at a level of at least 20 $\mu$ g/g of endosperm of said seed. In a particular embodiment said seed contains carotenoids at a level of at least 25 $\mu$ g/g of endosperm of said seed. In a particular embodiment said seed contains carotenoids at a level of at least 30 $\mu$ g/g of endosperm of said seed. In a particular embodiment said seed contains carotenoids at a level of at least 35 $\mu$ g/g of endosperm of said seed. In a particular embodiment said seed contains carotenoids at a level of at least 40 $\mu$ g/g of endosperm of said seed. In a particular embodiment said seed contains carotenoids at a level of at least 45 $\mu$ g/g of endosperm of said seed. In a particular embodiment said seed contains carotenoids at a level of at least 50 $\mu$ g/g of endosperm of said seed. In a particular embodiment said seed contains carotenoids at a level of at least 55 $\mu$ g/g of endosperm of said seed. In a particular embodiment said seed contains carotenoids at a level of at least 60 $\mu$ g/g of endosperm of said seed. In a particular embodiment said seed contains carotenoids at a level of at least 65 $\mu$ g/g of endosperm of said seed.

The present invention still further provides a method for increasing the carotenoid content of seeds comprising inserting into plant material a polynucleotide sequence as described above and regenerating a seed-containing plant from said material and identifying the seed which contains carotenoids amounting to at least 80% of the

- 17 -

carotenoid content of a seed which comprises a polynucleotide selected from the group depicted as SEQ ID NOS: 1; 2; 3; 4; 5 and 6. In a further embodiment, the seed of a plant regenerated from said material produce carotenoids amounting to at least 85% of the carotenoid content of a seed which comprises a polynucleotide selected from the group depicted as SEQ ID NOS: 1; 2; 3; 4; 5 and 6. In a still further embodiment, the seed of a plant regenerated from said material produce carotenoids amounting to at least 90% of the carotenoid content of a seed which comprises a polynucleotide selected from the group depicted as SEQ ID NOS: 1; 2; 3; 4; 5 and 6. In a still further embodiment, the seed of a plant regenerated from said material produce carotenoids amounting to at least 95% of the carotenoid content of a seed which comprises a polynucleotide selected from the group depicted as SEQ ID NOS: 1; 2; 3; 4; 5 and 6. In a still further embodiment, the seed of a plant regenerated from said material produce carotenoids amounting to at least 100% of the carotenoid content of a seed which comprises a polynucleotide selected from the group depicted as SEQ ID NOS: 1; 2; 3; 4; 5 and 6. In a particular embodiment the polynucleotide sequence provides for a percentage of the carotenoid content of seed as described above wherein the seed with which the comparison is made comprises the polynucleotide depicted as SEQ ID NO: 1. In a particular embodiment the polynucleotide sequence provides for a percentage of the carotenoid content of seed as described above wherein the seed with which the comparison is made comprises the polynucleotide depicted as SEQ ID NO: 2. In a particular embodiment the polynucleotide sequence provides for a percentage of the carotenoid content of seed as described above wherein the seed with which the comparison is made comprises the polynucleotide depicted as SEQ ID NO: 3. In a particular embodiment the polynucleotide sequence provides for a percentage of the carotenoid content of seed as described above wherein the seed with which the comparison is made comprises the polynucleotide depicted as SEQ ID NO: 4. In a particular embodiment the polynucleotide sequence provides for a percentage of the carotenoid content of seed as described above wherein the seed with which the comparison is made comprises the polynucleotide depicted as SEQ ID NO: 6.

The present invention still further provides a method for increasing the carotenoid content of seeds comprising inserting into plant material a polynucleotide comprising a sequence selected from the group depicted as SEQ ID NOS: 7; 8 and 9 and regenerating a seed containing plant from said material and identifying the seeds which contain carotenoids at levels greater than those of control like-seeds. In a particular embodiment

of the invention, the seeds obtained by said method contain at least a fifty fold increase in carotenoids when compared to a control like-seed.

The present invention still further provides a method as described above wherein said seed contains carotenoids at a level of at least 3µg/g of endosperm of said seed.

5       The present invention still further provides a method for increasing the carotenoid content of seeds comprising inserting into plant material a polynucleotide sequence as described above and regenerating a seed-containing plant from said material and identifying the seed which contains carotenoids amounting to at least 80% of the carotenoid content of a seed which comprises a polynucleotide selected from the group depicted as SEQ ID NOS: 7; 8 and 9.

10       The present invention still further provides a method as described above wherein the carotenoids which are increased are selected from the group consisting of: lycopene; alpha-carotene; lutein; beta-carotene; zeaxanthin; beta-cryptoxanthin; antheraxanthin; violaxanthin; and neoxanthin or a combination thereof. In a further embodiment, the carotenoids which are increased are selected from the group consisting of: lycopene;  
15       alpha-carotene; lutein; beta-carotene; zeaxanthin; beta-cryptoxanthin; or a combination thereof. In a still further embodiment, the carotenoids which are increased are selected from the group consisting of: alpha-carotene; lutein; beta-carotene; zeaxanthin; beta-cryptoxanthin; or a combination thereof. In a still further embodiment, the carotenoids  
20       which are increased include at least phytoene and beta-carotene. In a still further embodiment, the carotenoids which are increased include at least beta-carotene. In a still further embodiment, the carotenoid which is increased is beta-carotene. In a still further embodiment the carotenoids which are increased are coloured carotenoids.

25       The polynucleotide or polynucleotide sequence or vector as described above may be inserted into plant material by plant transformation techniques that are well known to the person skilled in the art. Such techniques include but are not limited to particle mediated biolistic transformation, *Agrobacterium*-mediated transformation, protoplast transformation (optionally in the presence of polyethylene glycols); sonication of plant tissues, cells or protoplasts in a medium comprising the polynucleotide or vector; micro-  
30       insertion of the polynucleotide or vector into totipotent plant material (optionally employing the known silicon carbide "whiskers" technique), electroporation and the like. In a particular embodiment of the invention rice plant material is transformed in accordance with the methods described in the Examples disclosed herein. In a particular

embodiment the *Agrobacterium* that is used is a strain that has been modified to reduce the possibility of recombination between sequences having a high degree of similarity within the TDNA region of the *Agrobacterium*. Furthermore, techniques and elements such as those referred to in WO99/01563, US6,265,638, US5,731,179 and US5,591,616  
5 may be employed as part of the transformation process.

The present invention still further provides seed obtained or obtainable by a method as described above. In a specific embodiment said seed are rice seed. In a further embodiment said seeds are maize seeds.

Throughout this specification the terms "seed" and "seeds" may be interchanged  
10 with the terms "grain" or "grains". In particular, the terms "seed" and "seeds" refers to edible seeds or seed parts, in particular, seed endosperm.

The present invention still further comprises a plant which comprises a seed according to the preceding paragraph.

The present invention further provides a plant or plant material which comprises a  
15 polynucleotide or a polynucleotide sequence or a vector as described above. In a particular embodiment said plant or plant material is a rice plant or rice plant material or a maize plant or maize plant material. In a still further embodiment the plant or plant material of the present invention is selected from the group consisting watermelon, melon, mango, soybean, cotton, tobacco, sugar beet, oilseed rape, canola, flax, sunflower,  
20 potato, tomato, alfalfa, lettuce, maize, wheat, sorghum, rye, bananas, barley, oat, turf grass, forage grass, sugar cane, pepper, pea, field bean, rice, pine, poplar, apple, peach, grape, strawberry, carrot, cabbage, onion, citrus, cereal or nut plants or any other horticultural crops. In a specific embodiment said plant is a rice plant.

The present invention still further provides a plant or seed according to the  
25 invention which further comprises a gene which gene encodes an enzyme which is capable of converting carotene to a retinoid. An example of such a gene is the gene encoding  $\beta$ -carotene dioxygenase as described in WO01/48162 and/or WO01/48163.

The present invention still further provides a plant according to the invention which further comprises a polynucleotide which provides for a trait selected from the  
30 group consisting of: insect resistance and/or tolerance; nematode resistance and/or tolerance; herbicide resistance and/or tolerance; improved resistance and/or tolerance to stress; a substance having pharmaceutical activity and/or any other desired agronomic trait.

- 20 -

The present invention still further provides a plant according to the invention which further comprises a polynucleotide which provides for a further enhancement of isoprenoid biosynthesis and/or carotenoid accumulation in the plant. In a particular embodiment said polynucleotide provides for a transcription factor which provides for a further enhancement of isoprenoid biosynthesis and/or carotenoid accumulation in the plant.

The present invention still further provides a plant according to the invention which further comprises a polynucleotide which provides for an increase in plastids within the plant. In a particular embodiment said polynucleotide comprises the PhyA gene from oats or arabidopsis which genes are well known to the person skilled in the art. In a further embodiment said polynucleotide comprises the Hp1 or Hp2 gene from tomato which are also well known to the person skilled in the art.

The present invention still further provides a molecular marker which marker is capable of identifying plant material which comprises a sequence selected from the group depicted as SEQ ID NOS: 1 to 9 wherein said molecular marker comprises at least about 25 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 1 to 9.

The present invention still further provides a method for identifying plant material according to the invention via the use, for example, of the polymerase chain reaction (PCR). Suitable primers may be designed using parameters well known to those skilled in the art and based on the sequences listed in the Sequence Listing. The person skilled in the art is well versed in nucleic acid extraction techniques, and once a test sample has been isolated, it can be analysed for the presence of the sequence according to the invention using techniques that are well known in the art. These include, but are not limited to, PCR, RAPIDS, RFLPs and AFLPs.

The present invention still further provides a kit which kit comprises a means for obtaining a test sample and a means for detecting the presence of the sequences of the invention within said test sample. Kits may also be generated that are suitable for testing the carotenoid content of a test sample and this may optionally be combined with the features of a kit as described in the preceding sentence.

In a further aspect of the present invention there is provided the use of a polynucleotide, polynucleotide sequence or a vector as described above in a method for the production of seeds containing increased carotenoids. In a particular embodiment the

present invention provides the use of a polynucleotide selected from the group depicted as SEQ ID NOS: 1; 2; 3; 4; 5 and 6 for the production of seeds which contain carotenoids at levels greater than those of a control like-seed.

5 In a further aspect of the present invention there is provided the use of a polynucleotide, polynucleotide sequence or a vector as described above in a method for the production of seeds containing increased carotenoids. In a particular embodiment the present invention provides the use of a polynucleotide selected from the group depicted as SEQ ID NOS: 7; 8 and 9 for the production of seeds which contain carotenoids at levels greater than those of a control like-seed.

10 In a further aspect of the invention there is provided the use of a polynucleotide, polynucleotide sequence or a vector as described above in a method of producing a plant which comprises said polynucleotide, said polynucleotide sequence or said vector.

In a further aspect of the invention there is provided the use of a polynucleotide selected from the group depicted as SEQ ID NOS: 1; 2; 3; 4; 5 and 6 in a method for the production of a plant comprising said polynucleotide.

15 In a further aspect of the invention there is provided the use of a polynucleotide selected from the group depicted as SEQ ID NOS: 7; 8 and 9 in a method for the production of a plant comprising said polynucleotide.

In a further aspect of the invention there is provided a method for increasing the carotenoid content of seeds comprising inserting into plant material (a) a first polynucleotide which comprises as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a nucleotide sequence derived from a bacterium which sequence encodes a carotene desaturase; and (iii) a transcription termination region; and (b) a second polynucleotide which comprises as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a nucleotide sequence encoding a phytoene synthase which sequence is derived from maize (*Zea sp.*) or rice (*Orzya sp.*); and (iii) a transcription termination region; and (c) regenerating a seed containing plant from said material and identifying the seeds which contain carotenoids at levels greater than those of control like-seeds.

30 In a further aspect of the invention there is provided a method for increasing the carotenoid content of seeds comprising inserting into plant material (a) a first polynucleotide which comprises as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a nucleotide sequence derived from a

- 22 -

bacterium which sequence encodes a carotene desaturase; and (iii) a transcription termination region; and (b) a second polynucleotide which comprises as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a nucleotide sequence encoding a phytoene synthase which sequence is derived from  
5 tomato (*Lycopersicon sp.*) or pepper (*Capsicum sp.*); or a bacterium; and (iii) a transcription termination region; and (c) regenerating a seed containing plant from said material and identifying the seeds which contain carotenoids at levels greater than those of control like-seeds.

In a particular embodiment, step (a) of the preceding paragraph is performed prior  
10 to step (b). In a further embodiment, step (b) of the preceding paragraph is performed prior to step (a). In a still further embodiment, the promoter, the sequence encoding said carotene desaturase, the sequence encoding phytoene synthase and the terminator region are derived from the sequences depicted in the Sequence Listing. In a still further  
15 embodiment said carotene desaturase is the CrtI gene from *Erwinia sp.* In a still further embodiment the carotene desaturase comprises or consists of the sequence depicted as SEQ ID NO: 18 or 19. In a further embodiment said phytoene synthase is derived from maize or rice. In a still further embodiment said phytoene synthase is from maize or rice. In a still further embodiment said phytoene synthase comprises or consists of a sequence selected from the group depicted as SEQ ID NOS: 10; 11; 12; and 13.

20 In a further aspect of the invention there is provided a method for increasing the carotenoid content of seeds comprising crossing (a) a first plant comprising a polynucleotide which comprises as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a nucleotide sequence derived from a bacterium which sequence encodes a carotene desaturase; and (iii) a transcription  
25 termination region; with (b) a further plant comprising a polynucleotide which comprises as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a nucleotide sequence encoding a phytoene synthase which sequence is derived from maize (*Zea sp.*) or rice (*Oryza sp.*); and (iii) a transcription termination region; and (c) harvesting seed from the female parent of the thus crossed plants; and (d)  
30 growing said seed to produce plants comprising further seeds and identifying said further seeds which contain carotenoids at levels greater than those of control like-seeds.

In a further aspect of the invention there is provided a method for increasing the carotenoid content of seeds comprising crossing (a) a first plant comprising a

- 23 -

polynucleotide which comprises as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a nucleotide sequence derived from a bacterium which sequence encodes a carotene desaturase; and (iii) a transcription termination region; with (b) a further plant comprising a polynucleotide which comprises  
5 as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a nucleotide sequence encoding a phytoene synthase which sequence is derived from tomato (*Lycopersicon sp.*) or pepper (*Capsicum sp.*); or a bacterium; and (iii) a transcription termination region; and (c) harvesting seed from the female parent of the thus crossed plants; and (d) growing said seed to produce plants comprising further  
10 seeds and identifying said further seeds which contain carotenoids at levels greater than those of control like-seeds.

In a still further embodiment, the promoter, the sequence encoding said carotene desaturase, the sequence encoding phytoene synthase and the terminator region are obtainable from the sequences depicted in the Sequence Listing. In a still further  
15 embodiment said carotene desaturase is the CrtI gene from *Erwinia sp.* In a still further embodiment the carotene desaturase comprises or consists of the sequence depicted as SEQ ID NO: 18 or 19. In a further embodiment said phytoene synthase is derived from maize or rice. In a still further embodiment said phytoene synthase is from maize or rice. In a still further embodiment said phytoene synthase comprises or consists of a sequence  
20 selected from the group depicted as SEQ ID NOS: 10; 11; 12 and 13.

In a further aspect of the invention there is provided a polynucleotide which comprises (a) a region which comprises as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a nucleotide sequence derived from a bacterium which sequence encodes a carotene desaturase or a nucleotide sequence  
25 which encodes a carotene desaturase derived from a plant selected from the group consisting of: tomato (*Lycopersicon sp.*); pepper (*Capsicum sp.*); maize (*Zea sp.*); rice (*Orzya sp.*); and (iii) a transcription termination region; and (b) a further region which comprises as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a nucleotide sequence encoding a phytoene synthase which  
30 sequence is derived from a bacterium, or from a plant selected from the group consisting of: tomato (*Lycopersicon sp.*); pepper (*Capsicum sp.*); maize (*Zea sp.*); rice (*Orzya sp.*); and (iii) a transcription termination region; and (c) a still further region which comprises as operably linked components (i) a promoter which provides for seed preferred



- 24 -

expression; and (ii) a nucleotide sequence encoding a zeta-carotene desaturase (ZDS) derived from a bacterium, or from a plant selected from the group consisting of: tomato (*Lycopersicon sp.*); pepper (*Capsicum sp.*); maize (*Zea sp.*); rice (*Orzya sp.*); and (iii) a transcription termination region. In a particular embodiment the carotene desaturase and phytoene synthase is derived from maize (*Zea sp.*) and and zeta-carotene desaturase (ZDS) is derived from pepper (*Capsicum sp.*). In a further embodiment the carotene desaturase and phytoene synthase is from maize (*Zea sp.*) and the zeta-carotene desaturase (ZDS) is from pepper (*Capsicum sp.*).

The polynucleotides as described above may be used to identify polynucleotide sequences providing for a like-function, based on the hybridisation conditions described above. These polynucleotides and polynucleotide sequences which comprise the zeta-carotene desaturase may also be used in methods for increasing the carotenoid content of seeds in a manner analogous to the methods described above.

In a further aspect of the invention there is provided a polynucleotide which comprises as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a nucleotide sequence derived from a bacterium which sequence encodes a carotene desaturase or a nucleotide sequence which encodes a carotene desaturase derived from a plant selected from the group consisting of: tomato (*Lycopersicon sp.*); pepper (*Capsicum sp.*); maize (*Zea sp.*); rice (*Orzya sp.*); and (iii) a nucleotide sequence encoding a phytoene synthase which sequence is derived from a plant selected from the group consisting of: tomato (*Lycopersicon sp.*); pepper (*Capsicum sp.*); maize (*Zea sp.*); rice (*Orzya sp.*) or a bacterium; and (iv) a nucleotide sequence encoding a zeta-carotene desaturase (ZDS) derived from a bacterium or from a plant selected from the group consisting of: tomato (*Lycopersicon sp.*); pepper (*Capsicum sp.*); maize (*Zea sp.*); rice (*Orzya sp.*); and (v) a transcription termination region.

Any of the regions described in this specification may be separated by a region which provides for a self-processing polypeptide which is capable of separating the proteins such as the self-processing polypeptide described in US5,846,767 or any similarly functioning element. Alternatively the regions may be separated by a sequence which acts as a target site for an external element which is capable of separating the protein sequences. Alternatively the polynucleotide may provide for a polyprotein which comprises a plurality of protein functions. In a further embodiment of the present

- 25 -

invention the proteins of the polyprotein may be arranged in tandem. The person skilled in the art will appreciate that when a polynucleotide is generated which encodes such a polyprotein, expression of such a polyprotein may be achieved via the use of a single promoter which promoter is described herein.

5 All of the polynucleotides and polynucleotide sequences described throughout this specification can be isolated and constructed using techniques that are well known to the person skilled in the art. For example, the polynucleotides can be synthesised using standard polynucleotide synthesisers. Such synthetic polynucleotides can be synthesised and then ligated to form the longer polynucleotides according to the invention. The  
10 polynucleotides and polynucleotide sequences can also be isolated from other constructs/vectors which contain said sequences and then be inserted into further constructs/vectors to produce the ones according to the invention. The sequences can also be isolated from libraries, for example cDNA and gDNA, using the sequence information provided in the Sequence Listing for the creation of suitable probes/primers  
15 for the purpose of identifying said sequences from said libraries. Once isolated, the sequences can be assembled to create the polynucleotides and polynucleotide sequences according to the invention. The sequences according to the invention can also be used to identify like-sequences in accordance with the hybridisation conditions described above. In identifying such like-sequences the person skilled in the art may wish to identify like-  
20 sequences of one or more of the component parts of the sequences depicted in the Sequence Listing and then subsequently assemble the like-sequences in a manner similar to the arrangement of the sequences depicted in the Sequence Listing. For example, it may be desired to modify the region encoding phytoene synthase gene only. Once the modified phytoene synthase encoding sequence had been identified, it could be used to  
25 replace the phytoene synthase encoding sequence in one of the sequences depicted in the Sequence Listing. Alternatively, the modified phytoene synthase may be used in the creation of a new sequence wherein all the component parts are the same as a sequence in the Sequence Listing save for the phytoene synthase. In a further example, all of the components may be modified and then each of the modified components is arranged in  
30 the same manner as the sequences of the Sequence Listing, for example, promoter-intron-target sequence-carotene desaturase-terminator- promoter-intron-(target sequence) - phytoene synthase-terminator. The degree of modification will affect the ability of the

-35-

tag

1263

&lt;210&gt; 14

&lt;211&gt; 420

&lt;212&gt; PRT

&lt;213&gt; Oryza sp.

&lt;400&gt; 14

Met Ala Ala Ile Thr Leu Leu Arg Ser Ala Ser Leu Pro Gly Leu Ser  
1 5 10 15

Asp Ala Leu Ala Arg Asp Ala Ala Ala Val Gln His Val Cys Ser Ser  
20 25 30

Tyr Leu Pro Asn Asn Lys Glu Lys Lys Arg Arg Trp Ile Leu Cys Ser  
35 40 45

Leu Lys Tyr Ala Cys Leu Gly Val Asp Pro Ala Pro Gly Glu Ile Ala  
50 55 60

Arg Thr Ser Pro Val Tyr Ser Ser Leu Thr Val Thr Pro Ala Gly Glu  
65 70 75 80

Ala Val Ile Ser Ser Glu Gln Lys Val Tyr Asp Val Val Leu Lys Gln  
85 90 95

Ala Ala Leu Leu Lys Arg His Leu Arg Pro Gln Pro His Thr Ile Pro  
100 105 110

Ile Val Pro Lys Asp Leu Asp Leu Pro Arg Asn Gly Leu Lys Gln Ala  
115 120 125

Tyr His Arg Cys Gly Glu Ile Cys Glu Glu Tyr Ala Lys Thr Phe Tyr  
130 135 140

Leu Gly Thr Met Leu Met Thr Glu Asp Arg Arg Arg Ala Ile Trp Ala  
145 150 155 160

Ile Tyr Val Trp Cys Arg Arg Thr Asp Glu Leu Val Asp Gly Pro Asn  
165 170 175

Ala Ser His Ile Thr Pro Ser Ala            Leu Asp Arg Trp Glu Lys Arg Leu  
                    180                          185                          190

Asp Asp Leu Phe Thr Gly Arg Pro Tyr Asp Met Leu Asp Ala Ala Leu  
195 200 205

Ser Asp Thr Ile Ser Lys Phe Pro Ile Asp Ile Gln Pro Phe Arg Asp  
210 215 220

Met Ile Glu Gly Met Arg Ser Asp Leu Arg Lys Thr Arg Tyr Lys Asn  
225 230 235 240

Phe Asp Glu Leu Tyr Met Tyr Cys Tyr Tyr Val Ala Gly Thr Val Gly  
245 250 255

Leu Met Ser Val Pro Val Met Gly Ile Ala Pro Glu Ser Lys Ala Thr  
260 265 270

Thr Glu Ser Val Tyr Ser Ala Ala Leu Ala Leu Gly Ile Ala Asn Gln  
275 280 285

Leu Thr Asn Ile Leu Arg Asp Val Gly Glu Asp Ala Arg Arg Gly Arg  
290 295 300

Ile Tyr Leu Pro Gln Asp Glu Leu Ala Glu Ala Gly Leu Ser Asp Glu  
305 310 315 320

Asp Ile Phe Asn Gly Val Val Thr Asn Lys Trp Arg Ser Phe Met Lys  
325 330 335

Arg Gln Ile Lys Arg Ala Arg Met Phe Phe Glu Glu Ala Glu Arg Gly  
340 345 350

Val Thr Glu Leu Ser Gln Ala Ser Arg Trp Pro Val Trp Ala Ser Leu  
355 360 365

Leu Leu Tyr Arg Gln Ile Leu Asp Glu Ile Glu Ala Asn Asp Tyr Asn  
370 375 380

Asn Phe Thr Lys Arg Ala Tyr Val Gly Lys Ala Lys Lys Leu Leu Ala  
385 390 395 400

Leu Pro Val Ala Tyr Gly Arg Ser Leu Leu Met Pro Tyr Ser Leu Arg  
405 410 415

-37-

Asn Ser Gln Lys  
420

&lt;210&gt; 15

&lt;211&gt; 1260

&lt;212&gt; DNA

&lt;213&gt; Capsicum annuum

&lt;400&gt; 15

atgtctgttg ccttgttatg ggttgtttct ccttgtagacg tctcaaacgg gacaggattc	60
ttggtatccg ttctgtaggg aaaccggatt tttgattcgt cggggcgtag gaatttggcg	120
tgcaatgaga gaatcaagag aggaggtgga aaacaaaggt ggagttttgg ttcttacttg	180
ggaggagcac aaactggaag tggacggaaa ttttctgtac gttctgctat cgtggctact	240
ccggctggag aaatgacgat gtcatacaga cggatggtat atgatgtggt tttgaggcag	300
gcagccttgg tgaagagaca gctgagatcg accgatgagt tagatgtgaa gaaggatata	360
cctattccgg ggactttggg cttggtgagt gaagcatatg ataggtgtag tgaagtatgt	420
gcagagtacg caaagacgtt ttacttagga acgatgctaa tgactccgga gagaagaaag	480
gctatctggg caatatacgt atggtgcagg agaacagacg aacttggtga tgggtccgaat	540
gcatcacaca ttactccggc ggccttagat aggtgggaag acaggctaga agatgttttc	600
agtggacggc catttgacat gctcgatgct gctttgtccg acacagtttc caaatttcca	660
gttgatatcc agccattcag agatatgatt gaaggaatgc gtatggactt gaggaagtca	720
agatacagaa actttgacga actataccta tattgttatt acgttgctgg tacggttggg	780
ttgatgagtg ttccaattat gggcatcgca cctgaatcaa aggcaacaac ggagagcgta	840
tataatgctg ctttggtctt ggggatcgca aatcagctga ccaacatact tagagatgtt	900
ggagaagatg ccagaagagg aagagtctat ttgcctcaag atgaattagc acaggcaggt	960
ctatccgacg aagacatatt tgctggaaga gtgaccgata aatggagaat cttcatgaag	1020
aaacaaattc agagggcaag aaagtctctt gacgaggcag agaaaggagt gaccgaattg	1080
agcgcagcta gtagatggcc tgtgttgga tctctgctgt tgtaccgcag gatactggac	1140
gagatcgaag ccaatgacta caacaacttc acaaagagag cttatgtgag caaaccaaag	1200
aagttgattg cattacctat tgcatacgca aaatctcttg tgccctctac aagaacatga	1260

&lt;210&gt; 16

-38-

&lt;211&gt; 1239

&lt;212&gt; DNA

<213> *Lycopersicon esculentum*

&lt;400&gt; 16

```
atgtctgttg ccttggtatg ggttggttct ccttgtagacg tctcaaattg gacaagtttc      60
atggaatcag tccgggaggg aaaccgtttt ttgattcat cgaggcatag gaatttggtg      120
tccaatgaga gaatcaatag aggtggtgga aagcaaacta ataatggacg gaaattttct      180
gtacggctctg ctattttggc tactccatct ggagaacgga cgatgacatc ggaacagatg      240
gtctatgatg tggttttgag gcaggcagcc ttggtgaaga ggcaactgag atctaccaat      300
gagttagaag tgaagccgga tatacctatt cgggggaatt tgggcttggt gagtgaagca      360
tatgataggt gtggtgaagt atgtgcagag tatgcaaaga cgtttaactt aggaactatg      420
ctaagtactc ccgagagaag aagggtatc tgggcaatat atgtatggtg cagaagaaca      480
gatgaacttg ttgatggccc aaacgcata tatattacc cggcagcctt agataggtgg      540
gaaaataggc tagaagatgt tttcaatggg cggccatttg acatgctcga tggtgctttg      600
tccgatacag tttctaactt tccagttgat attcagccat tcagagatat gattgaagga      660
atgcgtatgg acttgagaaa atcgagatac aaaaacttcg acgaactata cttttattgt      720
tattatgttg ctggtacggg tgggttgatg agtggtccaa ttatgggtat cgccctgaa      780
tcaaaggcaa caacagagag cgtatataat gctgctttgg ctctggggat cgcaaatcaa      840
ttaactaaca tactcagaga tgttgagaaa gatgccagaa gaggaagagt ctacttgctt      900
caagatgaat tagcacaggc aggtctatcc gatgaagata tatttgctgg aagggtgacc      960
gataaatgga gaatctttat gaagaaacaa atacataggg caagaaagtt ctttgatgag     1020
gcagagaaaag gcgtgacaga attgagctca gctagtagat tccctgtatg ggcatctttg     1080
gtcttgtacc gcaaaatact agatgagatt gaagccaatg actacaacaa cttcaciaag     1140
agagcatatg tgagcaaata aaagaagttg attgcattac ctattgcata tgcaaaatct     1200
cttgtgcctc ctacaaaaac tgcctctctt caaagataa                               1239
```

&lt;210&gt; 17

&lt;211&gt; 891

&lt;212&gt; DNA

<213> *Erwinia* sp.

-39-

<400> 17  
 atggcagttg gctcgaaaag ttttgcgaca gcctcaaagt tatttgatgc aaaaacccgg 60  
 cgcagcgtag tgatgctcta cgcctgggtg cgccattgtg acgatgttat tgacgatcag 120  
 acgctggggt ttcaggcccc gcagcctgcc ttacaaacgc ccgaacaacg tctgatgcaa 180  
 cttgagatga aaacgcgcca ggcctatgca ggatcgagcaga tgcacgaacc ggcgtttgag 240  
 gcttttcagg aagtggctat ggctcatgat atcgccccgg cttacgcgtt tgatcatctg 300  
 gaaggcttcg cgatggatgt acgcgaagcg caatacagcc aactggatga tacgctgcgc 360  
 tattgctatc acgttgcagg cgttgtcggc ttgatgatgg cgcaaatacat ggcgtgcgg 420  
 gataacgcca cgctggaccg cgcctgtgac cttggggtgg catttcagtt gaccaatatt 480  
 gctcgcgata ttgtggacga tgcgcatgag ggccgctgtt atctgccggc aagctggctg 540  
 gagcatgaag gtctgaacaa agagaattat gcggcacctg aaaaccgtca ggcgtgagc 600  
 cgtatcgccc gacgtttggg gcaggaagca gaaccttact atttgtctgc cacagccggc 660  
 ctggcagggg tgcacctgag ttccgcctgg gcaatcgcta cggcgaagca ggtttaccgg 720  
 aaaataggtg tcaaagttga acaggccggg cagcaagcct gggatcagcg gcagtcaacg 780  
 aacacgcccc aaaaattaac gctgctgctg gccgcctctg gtcaggccct tacttcccg 840  
 atgcgggctc atcctccccg cctgcgcgat ctctggcagc gcccgctcta g 891

&lt;210&gt; 18

&lt;211&gt; 1479

&lt;212&gt; DNA

<213> *Erwinia* sp.

<400> 18  
 atgaaaccaa ctacggtaat tgggtgcaggc ttcgggtggcc tggcactggc aattcgtcta 60  
 caagctgcgg ggatccccgt cttactgctt gaacaacgtg ataaacccgg cggtcgggct 120  
 tatgtctacg aggatcaggg gtttaccttt gatgcaggcc cgacggttat caccgatccc 180  
 agtgccattg aagaactggt tgcactggca ggaaaacagt taaaagagta tgtcgaactg 240  
 ctgccgggta cgccgtttta ccgcctgtgt tgggagtcag ggaaggtctt taattacgat 300  
 aacgatcaaa cccggctcga agcgcagatt cagcagttta atccccgca tgtcgaaggt 360  
 tatcgtcagt ttctggacta ttcacgcgcg gtgtttaaag aaggctatct gaagctcggg 420  
 actgtccctt ttttatcggt cagagacatg cttcgcgcgg cacctcaact ggcgaaactg 480

-40-

caggcatgga gaagcggtta cagtaagggt gccagttaca tcgaagatga acatctgcgc  
540

caggcggttt ctttccactc gctgttggtg ggcggcaatc ccttcgccac ctcattccatt 600  
tatacggtga tacacgcgct ggagcgtgag tggggcgtct ggtttccgcg tggcggcacc 660  
ggcgcattag ttcaggggat gataaagctg tttcaggatc tgggtggcga agtcgtgtta 720  
aacgccagag tcagccatat ggaaacgaca ggaaacaaga ttgaagccgt gcatttagag 780  
gacggtcgca ggttcttgac gcaagccgtc gcgtcaaatg cagatgtggt tcatacctat 840  
cgcgacctgt taagccagca cctgcccgcg gttaagcagt ccaacaaact gcagactaag 900  
cgcagtagta actctctgtt tgtgctctat tttggtttga atcaccatca tgatcagctc 960  
gcgcatcaca cggtttggtt cggcccgcgt taccgcgagc tgattgacga aatttttaat 1020  
catgatggcc tcgcagagga cttctcactt tatctgcacg cgcctgtgt caccgattcg 1080  
tcactggcgc ctgaagggtg cggcagttac tatgtgttg cgcgggtgcc gcatttaggc 1140  
accgcaacc tcgactggac ggttgagggg ccaaaactac gcgaccgtat ttttgcgta 1200  
cttgagcagc attacatgcc tggcttacgg agtcagctgg tcacgcaccg gatgtttacg 1260  
ccgtttgatt ttcgcgacca gcttaatgcc tatcatggct cagccttttc tgtggagccc 1320  
gttcttacc agagcgctg gtttcggccg cataaccgcg ataaaacat tactaatctc 1380  
tacctggctg gcgcaggcac gcatcccggc gcaggcatc ctggcgctcat cggctcggca 1440  
aaagcgacag caggtttgat gctggaggat ctgatttga 1479

&lt;210&gt; 19

&lt;211&gt; 1488

&lt;212&gt; DNA

&lt;213&gt; Erwinia sp.

&lt;400&gt; 19

atggcgccg ccaaaccaac tacggaatt ggtgcaggct tcggtggcct ggcactggca 60  
attcgtctac aagctgcggg gatccccgtc ttactgcttg aacaacgtga taaaccggc 120  
ggtcgggctt atgtctacga ggatcagggg ttacctttg atgcaggccc gacggttatc 180  
accgatccca gtgccattga agaactgttt gcaactggcag gaaaacagtt aaaagagtat 240  
gtcgaactgc tgccggttac gccgttttac cgcctgtgtt gggagtcagg gaaggtcttt 300  
aattacgata acgatcaaac ccggctcgaa gcgcagattc agcagtttaa tccccggat 360  
gtcgaagggt atcgtcagtt tctggactat tcacgcgcgg tgtttaaaga aggtatctg 420



-41-

aagctcggta ctgtcccttt tttatcggtc agagacatgc ttcgcgccgc acctcaactg 480  
 gcgaaactgc aggcattggag aagcggttac agtaagggtg ccagttacat cgaagatgaa 540  
 catctgcgcc aggcgttttc tttccactcg ctgttggtgg gcggcaatcc cttcgccacc 600  
 tcatccattt atacgttgat acacgcgctg gagcgtgagt ggggcgtctg gtttccgcgt 660  
 ggcggcaccg gcgcattagt tcaggggatg ataaagctgt ttcaggatct gggatggcgaa 720  
 gtctgtgttaa acgccagagt cagccatatg gaaacgacag gaaacaagat tgaagccgtg 780  
 catttagagg acggtcgcag gttcctgacg caagccgtcg cgtcaaactg agatgtgggt 840  
 catacctatc gcgacctgtt aagccagcac cctgccgcgg ttaagcagtc caacaaactg 900  
 cagactaagc gcatgagtaa ctctctgttt gtgctctatt ttggtttgaa tcaccatcat 960  
 gatcagctcg cgcattcacac gggttggttc ggcccgctt accgcgagct gattgacgaa 1020  
 atttttaatc atgatggcct cgcagaggac ttctcacttt atctgcacgc gccctgtgtc 1080  
 acggattcgt cactggcgcc tgaagggtgc ggcagttact atgtgttggt gccggtgcgc 1140  
 catttaggca ccgcgaacct cgactggacg gttgaggggc caaaactacg cgaccgtatt 1200  
 tttgcgtacc ttgagcagca ttacatgcct ggcttacgga gtcagctggt cagcaccgg 1260  
 atgtttacgc cgtttgattt tcgcgaccag cttaatgcct atcatggctc agcctttct 1320  
 gtggagcccg ttcttaccga gagcgccctg tttcgccgcg ataaccgcga taaaaccatt 1380  
 actaatctct acctggtcgg cgcaggcacg catcccggcg caggcattcc tggcgtcatc 1440  
 ggctcggcaa aagcgacagc aggtttgatg ctggaggatc tgatttga 1488

&lt;210&gt; 20

&lt;211&gt; 839

&lt;212&gt; DNA

<213> *Oryza* sp.

&lt;400&gt; 20

gtaaatcatg gtgtaggcaa cccaaataaa acaccaaact atgcacaagg cagtttggtg 60  
 tattctgtag tacagacaaa actaaaagta atgaaagaag atgtggtggt agaaaaggaa 120  
 acaatatcat gagtaatgtg tgagcattat gggaccacga aataaaaaga acattttgat 180  
 gagtcgtgta tcctcgatga gcctcaaaag ttctctcacc ccggataaga aacccttaag 240  
 caatgtgcaa agtttgcatt ctccactgac ataatgcaaa ataagatatc atcgatgaca 300  
 tagcaactca tgcattcatat catgcctctc tcaacctatt cattcctact catctacata 360

-42-

```

agtatcttca gctaaatggt agaacataaa cccataagtc acgtttgatg agtattaggc      420
gtgacacatg acaaatacaca gactcaagca agataaagca aaatgatgtg tacataaaac      480
tccagagcta tatgtcatat tgcaaaaaga ggagagctta taagacaagg catgactcac      540
aaaaattcat ttgcctttcg tgtcaaaaag aggagggctt tacattatcc atgtcatatt      600
gcaaaagaaa gagagaaaga acaacacaat gctgcgtcaa ttatacatat ctgtatgtcc      660
atcattattc atccaccttt cgtgtaccac acttcatata tcatgagtca cttcatgtct      720
ggacattaac aaactctatc ttaacattta gatgcaagag cctttatctc actataaatg      780
cacgatgatt tctcattggt tctcacaana agcattcagt tcattagtcc tacaacaac      839

```

&lt;210&gt; 21

&lt;211&gt; 642

&lt;212&gt; DNA

&lt;213&gt; Oryza sp.

```

<400> 21
aagcttgccg gcggaatacg gtggagatgg gttgggaacc ctggattcca aacacagccc      60
aagtctatcc aaaatgttta gacaagaaaa tacgtaacaa gttggtttac agaaatacga      120
attagatcaa tcctgcacta caagtagagt aaagtgggtga tttctcttaa atctctcgaa      180
tggtgattta agaattcagt gcaaaccaaa tccttgctat aatcaaatgt tcggtaccgc      240
atcaacggaa caataaaaag cgcctggcgt accataatth tgcattctt gttgaaatth      300
gtaatttaag atgcatgagg ccacacgacc ttaatgttca acgtgtcatg cattagttaa      360
ataatagctc aaaaaacgca acaaatatag ctagataacg gttgcaatcc ttaccaaact      420
aacgtataaa gtgagcgatg agtcatatca ttatctcccg cctgctaacc atcgtgtaca      480
ccatccgacg acaaaaatga caacttctag ggatgaacct ggacaagggt tagggtttag      540
ggatgaatct ggacaaatga ttgttcaggt tcatccctag atgttggttt ctctgacgg      600
gacggaggga gtatatgtga tggacacaaa agttactttc at                        642

```

&lt;210&gt; 22

&lt;211&gt; 190

&lt;212&gt; DNA

&lt;213&gt; SYNTHETIC - INTRON

-43-

<400> 22  
gtaaatttct agtttttctc cttcattttc ttggtttagga cccttttctc tttttatttt 60  
tttgagcttt gatctttctt taaactgac tattttttta ttgattgggt atcgtgtaaa 120  
tattacatag ctttaactga taatctgatt actttatttc gtgtgtcttt gatcatcttg 180  
atagttacag 190

&lt;210&gt; 23

&lt;211&gt; 171

&lt;212&gt; DNA

&lt;213&gt; TRANSIT PEPTIDE

<400> 23  
atggettcta tgatatactc ttccgctgtg acaacagtca gccgtgcctc tagggggcaa 60  
tccgcgcgag tggtccatt cggcggcctc aaatccatga ctggattccc agtgaagaag 120  
gtcaacactg acattacttc cattacaagc aatgggtggaa gagtaaagtg c 171

&lt;210&gt; 24

&lt;211&gt; 254

&lt;212&gt; DNA

&lt;213&gt; NOS TERMINATOR

<400> 24  
gatcgttcaa acatttggca ataaagtttc ttaagattga atcctgttgc cggctcttgcg 60  
atgattatca tataatttct gttgaattac gtttaagcatg taataattaa catgtaatgc 120  
atgacgttat ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac 180  
gcatagaaa acaaaatata gcgcgcaaac taggataaat tatcgcgcg cgtgtcatct 240  
atgttactag atcg 254

&lt;210&gt; 25

&lt;211&gt; 193

&lt;212&gt; DNA

&lt;213&gt; CAMV 35S TERMINATOR REGION

-44-

&lt;400&gt; 25

gctgaaatca ccagtctctc tctacaaatc tatctctctc tataataatg tgtgagtagt 60  
 tcccagataa ggggaattagg gttcttatag ggtttcgctc atgtgttgag catataagaa 120  
 acccttagta tgtatttgta tttgtaaaat acttctatca ataaaatttc taattcctaa 180  
 aacccaaaatc cag 193

&lt;210&gt; 26

&lt;211&gt; 238

&lt;212&gt; DNA

&lt;213&gt; POTATO PROTEINASE INHIBITOR GENE TERMINATOR REGION

&lt;400&gt; 26

ccctagactt gtccatcttc tggattggcc aacttaatta atgtatgaaa taaaaggatg 60  
 cacacatagt gacatgctaa tcaactataat gtgggcatca aagttgtgtg ttatgtgtaa 120  
 ttactaatta tctgaataag agaaagagat catccatatt tcttatccta aatgaatgtc 180  
 acgtgtcttt ataattcttt gatgaaccag atgcatttta ttaaccaatt ccatatac 238

&lt;210&gt; 27

&lt;211&gt; 2321

&lt;212&gt; DNA

&lt;213&gt; Lycopersicon esculentum

&lt;400&gt; 27

gggtttatct cgcaagtgtg gctatggtgg gacgtgtcaa attttggatt gtagccaaac 60  
 atgagatttg atttaaaggg aattggccaa atcacgaaa gcaggcatct tcatcataaa 120  
 ttagtttggt tatttataca gaattatacg cttttactag ttatagcatt cggtatcttt 180  
 ttctgggtaa ctgccaaacc accacaaatt tcaagtttcc atttaactct tcaacttcaa 240  
 cccaaccaa tttatttget taattgtgca gaaccactcc ctatatcttc taggtgcttt 300  
 cattcgttcc gagtaaaatg cctcaaattg gacttgtttc tgctgttaac ttgagagtec 360  
 aaggtagttc agcttatctt tggagctcga ggtcgtcttc tttgggaact gaaagtcgag 420  
 atggttgctt gcaaaggaat tcgttatggt ttgctggtag cgaatcaatg ggtcataagt 480  
 taaagattcg tactcccat gccacgacca gaagattggt taaggacttg gggcctttaa 540

-45-

aggtcgtatg cattgattat ccaagaccag agctggacaa tacagttaac tatttgagg  
 600  
 ctgcattttt atcatcaacg ttccgtgctt ctccgcgccc aactaaacca ttggagattg 660  
 ttattgctgg tgcaggtttg ggtggtttgt ctacagcaaa atatttgga gatgctggtc 720  
 acaaaccgat actgctggag gcaagggatg ttctaggtgg aaaggtagct gcatggaaag 780  
 atgatgatgg agattggtac gagactgggt tgcataatatt ctttggggct taccxaaata 840  
 ttcagaacct gtttgagaa ttagggatta acgatcgatt gcaatggaag gaacattcaa 900  
 tgatatttgc aatgccaagc aagccaggag aattcagccg ctttgatttc tccgaagctt 960  
 taccgctcc tttaaatgga attttagcca tcttaaagaa taacgaaatg cttacatggc 1020  
 cagagaaagt caaatttgca attggactct tgccagcaat gcttgaggag caatcttatg 1080  
 ttgaagctca agatgggata agtggttaagg actggatgag aaagcaagggt gtgccggaca 1140  
 ggggtgacaga tgagggtgttc attgctatgt caaaggcact caactttata aaccctgacg 1200  
 aactttcaat gcagtgcatt ttgatcgcat tgaacagggt tcttcaggag aaacatgggt 1260  
 caaaaatggc ctttttagat ggtaatctc ctgagagact ttgcatgccg attgttgaac 1320  
 acattgagtc aaaagggtggc caagtcagac tgaactcacg aataaaaaag attgagctga 1380  
 atgaggatgg aagtgtcaag agttttatac tgagtgcagg tagtgcaatc gagggagatg 1440  
 cttttgtgtt tgccgctcca gtggatattt tcaagcttct attgcctgaa gactggaaag 1500  
 agattccata tttccaaaag ttggagaagt tagtcggagt acctgtgata aatgtacata 1560  
 tatggtttga cagaaaactg aagaacacat atgatcattt gctcttcagc agaagctcac 1620  
 tgctcagtgt gtatgctgac atgtctgtta catgtaagga atattacaac cccaatcagt 1680  
 ctatgttgga attggttttt gcacctgcag aagagtggat atctcgcagc gactcagaaa 1740  
 ttattgatgc aacgatgaag gaactagcaa cgctttttcc tgatgaaatt tcagcagatc 1800  
 aaagcaaagc aaaaatattg aagtaccatg ttgtcaaac tccgaggctt gtttataaaa 1860  
 ctgtgccagg ttgtgaacct tgctgggctt taaaagatc cccaatagag gggttttatt 1920  
 tagccggtga ctacacgaaa cagaaatact tggcttcaat ggaaggcgct gtcttatcag 1980  
 gaaagctttg tgctcaagct attgtacagg attatgagtt acttgttgga cgtagccaaa 2040  
 agaagttgtc ggaagcaagc gtagtttagc tttgtggtta ttatttagct tctgtacact 2100  
 aaatttatga tgcaagaagc gttgtacaca acatatagaa gaagagtgcg aggtgaagca 2160  
 agtaggagaa atgttaggaa agctcctata caaaaggatg gcatgttgaa gattagcacc 2220  
 ttttaaatcc caagtttaaa tataaagcat attttatgta ccactttctt tatctggggg 2280  
 ttgtaatccc tttatatctt tatgcaatct ttacgttagt t 2321

-46-

&lt;210&gt; 28

&lt;211&gt; 1749

&lt;212&gt; DNA

<213> *Capsicum annuum*

&lt;400&gt; 28

```
atgccccaaa ttggacttgt ttctgctgtc aacttgagag tccaaggtaa ttcagcttat      60
ctttggagct cgaggctctt tttgggaact gatagtcaag atggttgctc gcaaaggaat      120
tcgttatgtt ttggtggtag tgactcaatg agtcataggt taaagattcg taatccccat      180
tccataacga gaagattggc taaggatttc cggcctttta aggttgtttg cattgattat      240
ccaaggccag agctagacaa tacagttaac tatttgaggg ctgcattctt atcatcatca      300
ttccgatctt ctccgcgccc aaccaaacca ctggagattg ttattgctgg tgcaggtttg      360
ggtggtttgt ctacagcaaa atatttgga gatgctggtc acaaaccaat actgctggag      420
gcaagggatg ttctaggttg aaaggtagct gcatggaaag atgatgatgg agattggtat      480
gagactgggt tgcacatatt ctttggggct taccxaaata tgcagaacct atttgagaa      540
ttagggataa atgatcgatt gcaatggaag gaacattcga tgatatttgc aatgccaac      600
aagccaggag aattcagccg ctttgatttc ccgaagctt tacctgctcc tttaaattga      660
attttgga tccaaagaa caatgaaatg cttacatggc cagaaaaagt caaatttgca      720
attggactct tgccagcaat gcttggtggg caatcttatg ttgaagctca agacgggata      780
agtgttaagg actggatgag aaaacaaggt gtgccggata gggtgacgga tgaggtgttc      840
atcgccatgt caaaggcact taacttcata aatcctgatg agctttcgat gcagtgcac      900
ttgatcgctg tgaacagatt tcttcaggag aaacatggtt caaaaatggc ctttttagat      960
ggtaatcctc ctgagagact ttgcatgccg attgttgaac atatcgagtc aaaaggtgga     1020
caagtcagac tgaactcacg aataaaaaag attgagctga atgaggatgg aagtgtcaag     1080
tgttttatac tgaacgatgg tagtacaatt gagggagatg cttttgtgtt tgcgactcca     1140
gtggatattt tcaagcttct tttgcctgaa gactggaaag agattccata tttccaaaag     1200
ttggagaagt tagttggagt acctgtgata aatgtccata tatggtttga cagaaaactg     1260
aagaacacat ctgataattt gctcttcagc agaagccac tgctcagtgt gtatgctgac     1320
atgtccgtca catgtaagga atattacgac cccaacaagt ccatgttggg attggtcttt     1380
gcgcctgcag aagagtgggt atctcgagct gactctgaaa ttattgatgc tacaatgaag     1440
```

-47-

gaactagcaa agctatattcc tgatgaaatt tcggcggatc agagcaaagc aaaaatattg 1500  
 aagtatcatg ttgtcaaaac tccaaggtct gtatataaaa ctgtgccagg ttgtgaaccc 1560  
 tgtcggctct tgcaaagatc ccctgtagag gggttttatt tagctgggtga ctacacgaaa 1620  
 cagaaatact tggcttcaat ggaaggtgct gtcttatcag gaaagctttg tgcacaagct 1680  
 attgtacagg attacgagtt acttggtggc cggagccaga ggaagttggc agaaacaagt 1740  
 gtagtttag 1749

&lt;210&gt; 29

&lt;211&gt; 2264

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;400&gt; 29

ctccaaatgc ggaggtctcg actcttctct ctctctccat ctttatcacc gccccacgta 60  
 cacacccaat tctctgcaac tgggtctccc cgcctccacg acaactgccc ccgtctcaag 120  
 tccgcgcgct ccattcttca gctctcctat cctccgcta gaatatcttc atcggatattt 180  
 taccaacctg gatcaattta ctacgatac tctgaagcgt atacatatgc catatgggaa 240  
 atgacttcat agctgtgggt tgtcttatgg ctcttgaat ttgcagtagt ctgcctgtac 300  
 ctattggctg aagcagagct gacccccact ttatcaagag ttgctcaacg atggacactg 360  
 gctgcctgtc atctatgaat attactggag ctagccagac aagatctttt gcggggcaac 420  
 ttctctctca gagatgtttt gcgagtagtc actatacaag ctttgccgtg aaaaaacttg 480  
 tctcaaggaa taaaggaagg agatcacacc gtagacatcc tgccttgacg gttgtctgca 540  
 aggattttcc aagacctcca ctagaaagca caataaacta tttggaagct ggacagctct 600  
 cttcattttt tagaaacagc gaacgcccc gtaagccgtt gcaggctcgtg gttgctgggt 660  
 caggattggc tggctctatca acagcgaagt atctggcaga tgctggccat aaaccatatt 720  
 tgcttgaggc aagagatgtt ttgggtggaa aggtagctgc ttggaaggat gaagatggag 780  
 attggtacga gactgggctt catatatattt ttggagctta tcccaacata cagaatctgt 840  
 ttggcgagct taggattgag gatcgtttgc agtggaaaga acactctatg atattcgcca 900  
 tgccaaacaa gccaggagaa ttcagccggt tcgatttccc agaaactttg ccagcaccta 960  
 taaatgggat atgggccata ttgagaaaca atgaaatgct tacttggccg gagaaggtga 1020  
 agtttgcaat cggacttctg ccagcaatgg ttggtggtca accttatggt gaagctcaag 1080

-48-

atggcctaac cgtttcagaa tggatgaaaa agcaggggtgt tcctgatcgg gtgaacgatg 1140  
 aggtttttat tgcaatgtcc aaggcactca atttcataaa tcctgatgag ctatctatgc 1200  
 agtgcatttt gattgctttg aaccgatttc ttcaggagaa gcatgggttct aaaatggcat 1260  
 tcttggatgg taatccgcct gaaaggctat gcatgcctat tgttgatcac attcgggtcta 1320  
 ggggtggaga ggtccgcctg aattctcgta ttaaaaagat agagctgaat cctgatggaa 1380  
 ctgtaaaaca cttcgcactt agtgatggaa ctcaaataac tggagatgct tatgtttgtg 1440  
 caacaccagt cgatatcttc aagcttcttg tacctcaaga gtggagtga attacttatt 1500  
 tcaagaaact ggagaagttg gtgggagttc ctgttatcaa tgttcatata tggtttgaca 1560  
 gaaaactgaa caacacatat gaccaccttc ttttcagcag gagttcactt ttaagtgtct 1620  
 atgcagacat gtcagtaacc tgcaaggaat actatgaccc aaaccgttca atgctggagt 1680  
 tggcttttgc tcctgcagac gaatggattg gtcgaagtga cactgaaatc atcgatgcaa 1740  
 ctatggaaga gctagccaag ttatttcctg atgaaattgc tgctgatcag agtaaagcaa 1800  
 agattcttaa gtatcatatt gtgaagacac cgagatcggg ttacaaaact gtcccaaact 1860  
 gtgagccttg ccggcctctc caaaggtcac ctatcgaagg tttctatcta gctggtgatt 1920  
 acacaaagca gaaatacctg gcttctatgg aagggtgcagt cctatccggg aagctttgtg 1980  
 cccagtccat agtgcaggat tatagcaggc tcgcactcag gagccagaaa agcctacaat 2040  
 caggagaagt tcccgcccca tcttagttgt agttggcttt agctatcgtc atccccactg 2100  
 ggtgctatct tatctcctat ttcaatggga acccacccaa tggatcatgtt ggagacaaca 2160  
 cctgttatgg tcctttgacc atctcgtggt gactgtagtt gatgtcatat tcggatatat 2220  
 atgtaaaagg acctgcatag caattgttag accttggaag aaaa 2264

&lt;210&gt; 30

&lt;211&gt; 2027

&lt;212&gt; DNA

<213> *Oryza* sp.

&lt;400&gt; 30

gtttatgaca gcatctgccg gatattttgc aggacaactt cctactcata ggtgcttcgc 60  
 aagtagcagc atccaagcac tgaaaggtag tcagcatgtg agctttggag tgaaatctct 120  
 tgtcttaagg aataaaggaa aaagattccg tcggaggctc ggtgctctac aggttggttg 180  
 ccaggacttt ccaagacctc cactagaaaa cacaataaac tttttggaag ctggacaact 240



-49-

atcctcattt ttcagaaaca gtgaacaacc cactaaacca ttacaggtcg tgattgctgg	300
agcaggatta gctggtttat caacggcaaa atatctggca gatgctggtc ataaacccat	360
attgcttgag gcaagggatg ttttgggtgg aaagatagct gcttgggaagg atgaagatgg	420
agattggtat gaaactgggc ttcatatctt ttttggagct tatcccaaca tacagaactt	480
gtttggcgag cttggtatta atgatcggtt gcaatggaag gaacactcca tgatatttgc	540
catgccaaac aagccaggag aatccagccg gtttgatttt cctgaaacat tgcctgcacc	600
cttaaatgga atatgggcca tactaagaaa caatgaaatg ctaacttggc cagagaaggt	660
gaagtttgct cttggacttt tgccagcaat ggttgggtggc caagcttatg ttgaagctca	720
agatggtttt actgtttctg agtggatgaa aaagcagggt gttcctgatc gagtgaacga	780
tgaagttttc attgcaatgt caaaggcact taatttcata aatcctgatg agttatccat	840
gcagtgcatt ctgattgctt taaaccgatt tcttcaggag aagcatgggt ctaagatggc	900
attcttggat ggtaatcctc ctgaaagggt atgcatgcct attgttgacc atgttcgctc	960
tttgggtggt gaggttcggc tgaattctcg tattcagaaa atagaactta atcctgatgg	1020
aacagtgaaa cactttgcac ttaccgatgg aactcaaata actggagatg cttatgtttt	1080
tgcaaacacca gttgatattc tgaagcttct tgtacctcaa gagtggaaag aaatatctta	1140
tttcaagaag ctggagaagt tgggtgggagt tctgtttata aatgttcata tatggtttga	1200
tagaaaactg aagaacacat atgaccacct tcttttcagc aggagttcac ttttaagtgt	1260
ttatgcggac atgtcagtaa cttgcaagga atactatgat ccaagccgtt caatgctgga	1320
gttggctctt gctcctgcag aggaatgggt tggacggagt gacactgaaa tcatcgaagc	1380
aactatgcaa gagctagcca agctatttcc tgatgaaatt gctgctgatc agagttaaagc	1440
aaagattctg aagtatcatg ttgtgaagac accaagatct gtttacaaga ctatcccgga	1500
ctgtgaacct tgccgacctc tgcaaagatc accgattgaa gggttctatc tagctgggtga	1560
ctacacaaag cagaaatatt tggcttcgat ggaggggtgca gttctatctg ggaagctttg	1620
tgctcagtct gtagtggagg attataaaat gctatctcgt aggagcctga aaagtctgca	1680
gtccgaagtt cctgttgctt cctagttgta gtcaggacta ttcccaatgg tgtgtgtgtc	1740
atcatccctt agtcagtttt tttctattta gtgggtgccc aactctccac caatttacac	1800
atgatggaac ttgaaagatg cctattttgg tcttatcata tttctgtaaa gttgatttgt	1860
gactgagagc tgatgccgat atgccacgct ggagaaaaag aacattatgt aaaacgacct	1920
gcatagtaat tcttagactt ttgcaaaagg caaaaggggt aaagcgacct tttttttcta	1980
tgtgaagggga ttaagagacc ttaaaaaaaaa aaaaaaaaaa aaaaaaa	2027

-50-

&lt;210&gt; 31

&lt;211&gt; 1931

&lt;212&gt; DNA

&lt;213&gt; Lycopersicon esculentum

&lt;400&gt; 31

```
ttttgtcttt ctttcttggt aacctttt cttgatattt aacaagaaaa gtttctttct      60
tttttttctt acctcataa ttgggtagag aacaattccc atggctactt cttcagctta      120
tctttcttgt cctgcaactt ctgctactgg aaagaaacat gttttcccaa atgggtcacc      180
tggattcttg gtttttggtg gtacctgtt gtccaaccgg ttagtgacct gaaagtcggt      240
tattcgggct gatttggatt ctatggtttc tgatatgagt accaacgctc caaaagggt      300
atttccacct gagcctgaac attatcgggg gccaaagctg aaagtagcta ttattggagc      360
tgggcttgca ggcattgca ctgctgtgga gctcttgat caaggacatg aggtggatat      420
atacgaatca aggactttta ttgggtggaa agtgggttct tttgttgata gacgtgggaa      480
ccacattgaa atgggactgc acgtgttctt tgggtgttat aataatctgt tccgtctgtt      540
gaaaaagggt ggtgctgaaa aaaatctgct agtgaaggag catactcaca ctttgtaaa      600
taaagggggt gaaatagggg aacttgattt ccgctttcca gttggagcac ccttacatgg      660
aattaatgca tttctgtcta ctaatcagtt aaagatttat gataaagcta gaaatgctgt      720
agctcttgcc cttagtccag tggcggggc tttagttgat ccggatggtg cattgcagca      780
gatacgcgat ctagataatg taagcttttc tgagtggttt ctgtctaaag gtgggacgcg      840
tgctagcatc cagaggatgt gggatcctgt tgcatatgct cttggattca ttgactgtga      900
taacatgagt gtcggtgta tgctcactat atttgcatta tttgccaca aaacagaggg      960
ttccctatta cgcattgta aaggttctcc tgacgtttat ttgagtggc caattaagaa     1020
gtacatcatg gacaaagggg gcaggttcca tctgaggtgg ggatgcagag aggtactcta     1080
tgagacgtcc tctgatggaa gcatgtatgt tagtgggctt gccatgtcaa aggcactca     1140
gaagaaaatt gtaaaagctg atgcatatgt ggctgcatgt gatgtccctg gaattaaaag     1200
attggttcct cagaagtgga gggaattgga attctttgac aacatttaca aattggtcgg     1260
agtgcctgtt gttaccgtac aactacgcta caatggctgg gttacagagt tgcaggactt     1320
ggagcgttcg aggcaattga agcgcgctgc aggtatggac aatctctctc atacgccaga     1380
tgcagatttc tcttgctttg cagatcttgc attggcatct ccagatgatt actacattga     1440
```

-51-

gggacaaggc tcattgcttc aatgtgtcct tacacctggg gacctttaca tgcctctatc 1500  
 aaatgatgaa atcattaaaa gagttacaaa gcaggttttg gcattatttc ctctgctcca 1560  
 aggtcttgag gttacctggg catcagtttt gaagatagga caatctttat atcgtgaagg 1620  
 acctggtaaa gacccattca gacctgatca gaagacgcca gtggaaaatt tctttcttgc 1680  
 tggctcatat acaaaacagg actacatcga tagcatggaa ggagcaactc tttcaggtag 1740  
 gcaagcttct gcatacatat gtaatgttgg agagcagctg atggcggttc gtaaaaagat 1800  
 cactgctgct gagttgaatg acatctctaa aggtgtgtcc ctatctgatg agttgagtct 1860  
 tgtctgatga cagactgcaa atcatccaaa tacaactcag ttaggcacgc cacaaggaag 1920  
 aattcttcta a 1931

&lt;210&gt; 32

&lt;211&gt; 1982

&lt;212&gt; DNA

&lt;213&gt; Capsicum annuum

&lt;400&gt; 32

cacaattcta tggctacttg ttcagcttat ctttgttgc ctgccacttc tgcttcttta 60  
 aagaaacgtg tttttccaga tgggtccgct ggattcttgt tttttggtgg tcgtcgtttg 120  
 tcgaaccggt tagtgacccc aaagtctgtc atccgagctg atttgaactc catggctctc 180  
 gacatgagta ccaacgctcc aaaagggcta tttccacctg aacctgaaca ttatcggggg 240  
 ccaaagctga aagtagctat tattggagct ggccttgacg gcattgtcgc tgctgtggag 300  
 ctcttggaac aaggacatga ggtggatata tatgaatcaa ggaccttcac tgggtgggaaa 360  
 gtgggttctt ttgttgataa acgtgggaac cacattgaaa tgggactgca cgtgttcttt 420  
 ggttgctata ataacttatt ccgtctgatg aaaaagggtg gtgctgaaaa aaatctgcta 480  
 gtgaaggagc atactcacac atttgtaaat aaagggggtg aaatagggga gcttgatttc 540  
 cgctttccag ttggagcgcc cttacatgga attaattgat ttttgtctac taatcaacta 600  
 aagacttatg ataaagctag aaatgctgta gctcttgccc ttagtccagt ggtgcgggct 660  
 ttagttgatc cagatggcgc attgcagcag atacgtgatc tagatagtgt aagcttttct 720  
 gattggttta tgtctaaagg agggacgcgc gctagcatcc agaggatgtg ggatcctgtt 780  
 gcatatgctc ttggattcat tgactgtgac aatatcagtg ctgggtgtat gctcactata 840  
 tttgcattat ttgccactaa aacggaggct tccctactgc gcattgctaa aggttctcct 900

-52-

gacgtttatt tgagtgggcc aattaagaag tacatcatag acaagggggg aaggttccat 960  
 ctgaggtggg gatgcagaga ggtactctac gagacatcct ctgatggaag catgtatggt 1020  
 agcgggcttg ccatgtcaaa ggccactcag aagaaaattg taaaagctga tgcctatggt 1080  
 gccgcatgtg tagtacctgg aattaaaaga ttagtacctc agaagtggag ggaattggaa 1140  
 ttctttggca acatttaca actgattgga gtgcctgttg ttactgtgca actacgatac 1200  
 aatggctggg ttacggaggt gcaggacttg gagcggtcaa ggcaatcaaa gcgctgctaca 1260  
 ggtttggaca atctcctgta cagccagat gcagatttct cttgttttgc agaccttgca 1320  
 ttggcatctc cagaagatta ttacattgag ggacaaggct cgttgcttca atgtgtcctt 1380  
 acgcctggcg acccttacat gcctctacca aatgaagaaa tcataagaag agtgtcaaag 1440  
 caggttttgg cgttatttcc ttcttcccaa ggtcttgagg taacctgggc atcagttgtg 1500  
 aagattgggc aatccttata tcgtgaagga cctggtaaag acccgttcag acctgatcaa 1560  
 aagacgccag tggaaaattt ctttcttgct ggctcatata caaacagga ctacatcgat 1620  
 agtatggaag gggcaactct ttcaggcaga caagcttctg catacatatg tgatgctgga 1680  
 gagcagctgt tggcgtgctg aaaaaagatt gctgctgctg agttaaacga gatctctaaa 1740  
 ggtgtatcgc tatcgatga gttgagtctt gtctgatgac tgcaaatcat tcagaaatat 1800  
 aattcagtta ggcagtgc ataggaagaat tcttctaaat ttttgagtct cacaattatg 1860  
 gaaatcaaaa tatgttttaa aaatgttgta tgtatgtaat attagtaa atctcatagt 1920  
 atgtatgtat ctattctgcc acgcttcagt ttagtgaaat ggaacttatt gctgcatcaa 1980  
 tc 1982

&lt;210&gt; 33

&lt;211&gt; 2265

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;400&gt; 33

ccctgccacg acgcccgcga caaatccctg cgcgacggca tcttcgcctc ccacccctc 60  
 ccagcttccc ctccactcc ggccctcaca caaattgccc ctcttcttct cctcctcttt 120  
 acacgctgcc gaccacggct gccgccaacc acccgcccca cccgtccacc gctgccgagt 180  
 gctagccatt tggagctgcc gcgcatggc gtccgtggcc gccaccacca cgctggcacc 240  
 ggcactcgcc ccgcgccggg cgcggccagg gactgggctc gtgccgccgc gccgggcctc 300

-53-

ggccgctcgct gctcgctcga ccgtaacgct tccgacatgg cgtcaacgct cccaaagggt 360  
attccccaccc gagccagagc actacagggg cccgaagctc aagggtggcca tcataggggc 420  
aggccttgcg ggcattgtcca ccgctgttga gctcttggac caggggccatg aggttgattt 480  
gtacgagtcc cgtccgttta tcgggtggcaa ggttggctcc tttgttgaca ggcaaggaaa 540  
ccatatacgag atggggctgc atgtgttctt cgggtgctac agcaatctct tccgcctcat 600  
gaagaagggtt ggcgctgata ataattctgct ggtgaaggaa catacccata cttttgtaaa 660  
taaagggggc acgattgggtg aacttgattt tcgggtcccg gtgggagctc cgttacatgg 720  
cattcaagca ttcctaagaa ctaatcagct caaggtttat gataaagcaa gaaatgcagt 780  
tgctcttgcc cttagtccag ttgttcgggc tctggttgat cctgatgggtg cattgcagca 840  
agtgcggggc ttggatgata taagtttcag tgattgggtc atgtccaaag ggggtactcg 900  
ggagagtatc acaagaatgt gggatcctgt tcgttacgct ttgggtttca ttgactgtga 960  
taatatacgt gcacgttgca tgcctactat tttcaccttg tttgccacaa agacagagggc 1020  
atccctgtta cgcattgtta agggttcacc tgatgtttac ttaagtggtc caataaagaa 1080  
gtatataaca gacaggggtg gtaggtttca cttaagggtg ggatgcagag aggttctcta 1140  
tgagaagtca cctgatggag agacctatgt taagggcctt ctactacca aggctacaag 1200  
tagagagata atcaaagctg atgcatacgt cgcagcctgt gatgttccag gtatcaaaaag 1260  
attacttcca tcagaatgga gggagtggga aatgtttgac aatatctaca agttagatgg 1320  
tgtccctgtt gtcactgtcc agctccgcta caacggatgg gtcactgaac ttcaagattt 1380  
ggagaaatca agacaactgc aaagggcggg tgggttgat aacctttgt acacggcgga 1440  
tgcagacttt tctgtttttt cggaccttgc tctctcatct cctgctgatt actacattga 1500  
agggcaagggt tccctgatcc aagctgtgct gactcctgga gatccataca tgccattgcc 1560  
aaacgaggag atcattagta aggttcaaaa gcaggttgta gaactgttcc catcttcccg 1620  
gggcttagaa gttacatggt ccagtgtggt aaagatcgga caatcgctgt accgtgagggc 1680  
tcttgaaac gaccttca ggcctgatca gaagacgccc gttaaaaact tcttctcttc 1740  
tgatcttac acgaaacagg actacatcga cagcatggaa ggagcaactc tctccggcag 1800  
gcgaacgtcg gcctacatct gcggtgccgg ggaggagctg ctggccctcc gaaagaagct 1860  
actcatcgac gacggcgaga aggcgctggg gaacgttcaa gtctgcagg ctagctgaac 1920  
aaccctctct gcaactgcaga gaagcttgga tctttccaac cacacataca tgctggaatg 1980  
gacaaaccaa ccaaccattg tcttttctcg cttcaggggtg ctggcgattc ccgcagcaac 2040  
ctctgtgta tcgtatccaa tttgagcatt agatctgccc cccccctg caggcggttc 2100

-54-

tttcctatcc ctgatccgag aagcaggggtg tagtctaggt ggctggcata cgggattaca 2160  
tcaggcagtg tgtaagttca gctggaactc gattggtaat tgggatggat gattgatgat 2220  
atatatatag cacacactgt tcttgctgtc tgcaaaaaaa aaaaa 2265

&lt;210&gt; 34

&lt;211&gt; 1632

&lt;212&gt; DNA

&lt;213&gt; Oryza sp.

&lt;400&gt; 34

cccacgcgtc cgccacgcg tccgattgg tgaacttgat ttccggttc ctgtgggagc 60  
tccgttacat ggtatccaag cattcctacg aactaaccaa ctcaagggtt atgataaagc 120  
aagaaatgcc gttgctcttg ctctaagccc agttgttcga gctcttggtg atccagatgg 180  
tgcattgcag caagtacggg atttgatga tgtaagttc agcgattggt tcttgctgaa 240  
agtggttact cgagagagca tcacaaggat gtgggacct gttgcctatg ctcttggttt 300  
cattgactgt gataatatca gtgcacgttg catgcttacc attttcactc tgtttgccac 360  
aaaaacagag gcatctttat tacgcatgct aaagggttca cctgatgttt atctgagtgg 420  
tccaataaag aagtacataa cagacagggg tggtaggttt cacctgaggt ggggatgtag 480  
ggagggttctc tatgataagt cacctgatgg ggaaacctat gttaaaggcc ttctcctatc 540  
caaggctaca agtagagaga taatcaaagc agatgcata gtcgcagctt gtgatgtccc 600  
ggggatcaaa agacttttac cttctgaatg gaggcaatgg gatacatttg acaacatcta 660  
caagttagat ggtgttcttg tagtcacagt acagcttcgt tataatggat gggttacaga 720  
acttcaagat ttggagaaat caagacaact gaaaaaggca gttggcttgg ataattctct 780  
ctacactcca gatgcagatt ttcatgttt ttcagacctt gcactttcat ctctgctga 840  
ctactacatt gaaggacaag gttccttgat ccaagctgtg ctaaccctg gcgacctta 900  
catgccattg ccgaatgagg agataattag caaggttcaa aagcaggtct tagaattgtt 960  
cccgatca caaggttgg aacttacatg gtcgagtgtg gtgaaaatcg gtcaatcatt 1020  
gtaccgcgag tcaccaggaa atgatccatt tagacctgat caaaagacac cagttaaaaa 1080  
cttcttcttg tctggctctt acacaaaaca ggactacatt gacagcatgg aaggggcaac 1140  
tctctcaggc aggagaaccg cggcctacat ctgtggtgca ggagaggagc tgcttcgccc 1200  
tccgaaagaa gtcattgtc gacgacagcg gagaaggcca gggtaagggt cgacggccct 1260

-55-

tcagacaagc tgagcttcct caaatgacac atgctggagt gagtggattg ctatgcccaa 1320  
aacaaaaaca gcttcctggg tgtagtaggc gatttccgca gcgactctca tgtaaatacct 1380  
acttgattga gcatttaggt ccaatctgct gctgcccttt ttgccttgac acgatcggtc 1440  
gttcgccccgt caatgggtgtg ttcttcgtta ttgtgaattt gtgattggga accaaaggtg 1500  
gcatacgga ttacatcagg cagcgtgtgt tttgttcage ttaaccgatc attgaacca 1560  
ttgatgatga tgatgatgtt tatatagtgc acacatcact taataaaaaa aaaaaaaaaa 1620  
aaaaaaaaaa aa 1632

&lt;210&gt; 35

&lt;211&gt; 40

&lt;212&gt; DNA

&lt;213&gt; PRIMER

&lt;400&gt; 35

cgctggcctg catggcccta cttctggcta tttctcagtg

40

&lt;210&gt; 36

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; PRIMER

&lt;400&gt; 36

ctgtccatgg cggccatcac gctcct

26

&lt;210&gt; 37

&lt;211&gt; 40

&lt;212&gt; DNA

&lt;213&gt; PRIMER

&lt;400&gt; 37

cgatggcctg catggcccta ggtctggcca tttctcaatg

40

&lt;210&gt; 38

-56-

<211> 32

<212> DNA

<213> PRIMER

<400> 38

taggataaga tagcaaatcc atggccatca ta

32



-5-

<400> 2  
 gttaatcatg gtgtaggcaa cccaaataaa acacaaaaat atgcacaagg cagtttggtg 60  
 tattctgtag tacagacaaa actaaaagta atgaaagaag atgtggtggt agaaaaggaa 120  
 acaatatcat gagtaatgtg tgagcattat gggaccacga aataaaaaga acattttgat 180  
 gagtcgtgta tctctgatga gcctcaaaag ttctctcacc ccggataaga aacccttaag 240  
 caatgtgcaa agtttgcatt ctccactgac ataatgcaaa ataagatata atcgatgaca 300  
 tagcaactca tgcatacatat catgcctctc tcaacctatt cattcctact catctacata 360  
 agtatcttca gctaaatggt agaacataaa cccataagtc acgtttgatg agtattagge 420  
 gtgacacatg acaaatcaca gactcaagca agataaagca aaatgatgtg tacataaaac 480  
 tccagagcta tatgtcatat tgcaaaaaga ggagagctta taagacaagg catgactcac 540  
 aaaaattcat ttgcctttcg tgtcaaaaag aggagggcct tacattatcc atgtcatatt 600  
 gcaaaagaaa gagagaaaga acaacacaat gctgctgcaa ttatacatat ctgtatgtcc 660  
 atcattatcc atccaccttt cgtgtaccac acttcatata tcatgagtca cttcatgtct 720  
 ggacattaac aaactctatc ttaacattta gatgcaagag cctttatctc actataaatg 780  
 cacgatgatt tctcattggt tctcacaaaa agcattcagt tcattagtc tacaacaacg 840  
 aattcggtt cccgggtaca gggtaaatct ctagtttttc tcttcattt tcttggttag 900  
 gaccttttc tctttttatt tttttgagct ttgatctttc tttaaactga tctatttttt 960  
 aattgattgg ttatcgtgta aatattacat agcttttaact gataatctga ttactttatt 1020  
 tcgtgtgtct ttgatcatct tgatagttac agaaccgtcg actctagaga agccatttaa 1080  
 atcgccgcca ccatggcttc tatgatatac tcttcgctg tgacaacagt cagccgtgcc 1140  
 tctagggggc aatccgccc agtggtcca ttggcgggc tcaaatecat gactggattc 1200  
 ccagtgaaga aggtcaacac tgacattact tccattacaa gcaatggtgg aagagtaaag 1260  
 tgcataaac caactacgtt aattggtgca ggcttcggtg gcctggcact ggcaattcgt 1320  
 ctacaagctg cggggatccc cgtcttactg cttgaacaac gtgataaacc cggcggtcgg 1380  
 gcttatgtct acgaggatca ggggtttacc ttgatgcag gcccgacggt tatcaccgat 1440  
 cccagtcca ttgaagaact gtttgactg gcaggaaaac agttaaaga gtatgtcgaa 1500  
 ctgctgccgg ttacgccgtt ttaccgctg tgttgggagt cagggaagg ctttaattac 1560  
 gataacgatc aaaccggct cgaagcgag attcagcagt ttaatcccc cgatgtcgaa 1620  
 gggtatcgtc agtttctgga ctattcacgc gcggtgttta aagaaggcta tctgaagctc 1680  
 ggtactgtcc cttttttatc gttcagagac atgcttcgag ccgcacctca actggcgaaa 1740

-6-

ctgcaggcat ggagaagcgt ttacagtaag gttgccagtt acatcgaaga tgaacatctg  
 1800

cgccaggcgt tttctttcca ctgcgtgttg gtgggcggca atcccttcgc cacctcatcc 1860  
 atttatacgt tgatacacgc gctggagcgt gagtggggcg tctggtttcc gcgtggcggc 1920  
 accggcgcat tagttcaggg gatgataaag ctgtttcagg atctgggtgg cgaagtcgtg 1980  
 ttaaagccca gagtcagcca tatggaaacg acaggaaaca agattgaagc cgtgcattta 2040  
 gaggacggtc gcaggttctt gacgcaagcc gtcgcgtcaa atgcagatgt ggttcatacc 2100  
 tategcgacc tgtaagcca gcacctgcc gcggttaagc agtccaacaa actgcagact 2160  
 aagcgcatga gtaactctct gtttgtgctc tattttggtt tgaatcacca tcatgatcag 2220  
 ctgcgcgcat acacggtttg tttcgccccg cgttaccgcg agctgattga cgaaattttt 2280  
 aatcatgatg gcctcgcaga ggactttctc ctttatctgc acgcgccctg tgtcacggat 2340  
 tcgtcactgg cgctgaagg ttgcggcagt tactatgtgt tggcgccggg gccgcattta 2400  
 ggcaccgcga acctcgactg gacggttgag gggccaaaac tacgcgaccg tatttttgcg 2460  
 taccttgagc agcattacat gcctggctta cggagtcagc tggtcacgca ccggatgttt 2520  
 acgccgtttg attttcgcga ccagcttaat gcctatcatg gctcagcctt ttctgtggag 2580  
 cccgttctta ccagagcgc ctggtttcgg ccgcataacc gcgataaaac cattactaat 2640  
 ctctacctgg tcggcgccagg cacgcacccc ggcgcaggca ttcttggcgt catcggtcgc 2700  
 gcaaaagcga cagcaggttt gatgctggag gatctgattt gaggccatgc aggccgatcc 2760  
 ccgatcggtc aaacatttgg caataaagtt tcttaagatt gaatcctgtt gccgggtctt 2820  
 cgatgattat catataattt ctgttgaaatt acgttaagca tgtaataatt aacatgtaat 2880  
 gcatgacgtt atttatgaga tgggttttta tgattagagt cccgcaatta tacatttaat 2940  
 acgcgataga aaacaaaata tagcgcgcaa actaggataa attatcgcgc gcggtgtcat 3000  
 ctatgttact agatcgggcc ttaataagct tgtaatacat ggtgtaggca acccaaataa 3060  
 aacacaaaaa tatgcacaag gcagtttgtt gtattctgta gtacagacaa aactaaaagt 3120  
 aatgaaagaa gatgtggtgt tagaaaagga aacaatatca tgagtaatgt gtgagcatta 3180  
 tgggaccacg aaataaaaag aacattttga tgagtcgtgt atcctcgatg agcctcaaaa 3240  
 gttctctcac cccggataag aaacccttaa gcaatgtgca aagtttgcat tctccactga 3300  
 cataatgcaa aataagatat catcgatgac atagcaactc atgcatcata tcatgcctct 3360  
 ctcaacctat tcatctctac tcatctacat aagtatcttc agctaaatgt tagaacataa 3420  
 acccataagt cacgtttgat gagtattagg cgtgacacat gacaaatcac agactcaagc 3480  
 aagataaagc aaaatgatgt gtacataaaa ctccagagct atatgtcata ttgcaaaaag 3540

-7-

aggagagctt ataagacaag gcatgactca caaaaattca ttgaccttc gtgtcaaaaa 3600  
gaggagggct ttacattatc catgtcatat tgcaaaagaa agagagaaag aacaacacaa 3660  
tgctgcgtca attatacata tctgtatgtc catcattatt catccacctt tcgtgtacca 3720  
cacttcatat atcatgagtc acttcatgtc tggacattaa caaactctat cttaacattt 3780  
agatgcaaga gcctttatct cactataaat gcacgatgat ttctcattgt ttctcacaaa 3840  
aagcattcag ttcattagtc ctacaacaac gaattcggct tcccgggtac agggtaaatt 3900  
tctagttttt ctcttctatt ttcttgggta ggacctttt ctctttttat ttttttgagc 3960  
tttgatcttt ctttaaactg atctattttt taattgattg gttatcgtgt aaatattaca 4020  
tagctttaac tgataatctg attactttat ttcgtgtgtc tttgatcatc ttgatagtta 4080  
cagaaccgtc gactctagag aagccattta aatcgccgcc accatggcca tcatactcgt 4140  
acgagcagcg tcgccggggc tctccggcgc cgacagcatc agccaccagg ggactctcca 4200  
gtgtctccacc ctgtcaaga cgaagaggcc ggcggcgcgc cggtggatgc cctgctcgtc 4260  
ccttggcctc caccctgagg aggctggccg tccctccccc gccgtctact ccagcctcgc 4320  
cgtcaaccgg gcgggagagg ccgtcgtctc gtccgagcag aaggctctag acgtcgtgct 4380  
caagcaggcc gcattgctca aacgccagct gcgcacgccg gtctctgacg ccaggcccca 4440  
ggacatggac atgccacgca acgggctcaa ggaagcctac gaccgctgcg gcgagatctg 4500  
tgaggagtat gccaaagcgt ttacctcgg aactatgttg atgacagagg agcggcgcgcg 4560  
cgccatatgg gccatctatg tgtggtgtag gaggacagat gagctttag atggggccaaa 4620  
cgccaactac attacacca cagctttgga ccggtgggag aagagacttg aggatctgtt 4680  
cacgggacgt ccttacgaca tgcttgatgc cgctctctct gataccatct caaggttccc 4740  
catagacatt cagccattca gggacatgat tgaagggatg aggagtgatc ttaggaagac 4800  
aaggтатаac aacttcgacg agctctacat gtactgctac tatgttgctg gaactgtcgg 4860  
gttaatgagc gtaccagtga tgggcatcgc atccgagtct aaagcaacaa ctgaaagcgt 4920  
gtacagtgtc gccttggctc tcggaattgc gaaccaactc acgaacatac tccgggatgt 4980  
tgagaggat gctagacgag gaaggatata ttaccacaa gatgagcttg cacaggcagg 5040  
gctctctgat gaggacatct tcaaaggggt cgtcacgaac cggaggagaa acttcatgaa 5100  
gaggcagatc aagagggcca ggatgttttt tgaggaggca gagagagggg taactgagct 5160  
ctcacaggct agcagatggc cagtatgggc ttccctgttg ttgtacaggc agatcctgga 5220  
tgagatcgaa gccaacgact acaacaactt cacgaagagg gcgtatgttg gtaaagggaa 5280  
gaagttgcta gcacttctg tggcatatgg aaaatcgcta ctgctcccat gttcattgag 5340

-8-

aaatggccag acctagggcc atgcaggccg atccccgatc gttcaaacat ttggcaataa 5400  
agtttcttaa gattgaatcc tgttgccggg cttgcgatga ttatcatata atttctgttg 5460  
aattacgtta agcatgtaat aattaacatg taatgcatga cgttatttat gagatggggt 5520  
tttatgatta gagtcccgcga attatacatt taatacgcga tagaaaacaa aatatagcgc 5580  
gcaaactagg ataaattatc gcgcgcgggtg tcattctatgt tactagatcg 5630

&lt;210&gt; 3

&lt;211&gt; 5180

&lt;212&gt; DNA

&lt;213&gt; SYNTHETIC - 12422

&lt;400&gt; 3

gttaatcatg gtgtaggcaa cccaaataaa acacccaaat atgcacaagg cagtttgttg 60  
tattctgtag tacagacaaa actaaaagta atgaaagaag atgtggtgtt agaaaaggaa 120  
acaatatcat gagtaatgtg tgagcattat gggaccacga aataaaaaga acattttgat 180  
gagtcgtgta tcctcgatga gcctcaaaag ttctctcacc ccggataaga aacccttaag 240  
caatgtgcaa agtttgcatt ctccactgac ataatgcaa ataagatata atcgatgaca 300  
tagcaactca tgcacatata catgcctctc tcaacctatt cattcctact catctacata 360  
agtatcttca gctaaatgtt agaacataaa ccataagtc acgtttgatg agtattaggc 420  
gtgacacatg acaaatcaca gactcaagca agataaagca aaatgatgtg tacataaaac 480  
tccagagcta tatgtcatat tgcaaaaaga ggagagctta taagacaagg catgactcac 540  
aaaaattcat ttgcctttcg tgtcaaaaag aggagggtt tacattatcc atgtcatatt 600  
gcaaaagaaa gagagaaaga acaacacaat gctgcgtcaa ttatacatat ctgtatgtcc 660  
atcattatcc atccaccttt cgtgtaccac acttcatata tcatgagtca cttcatgtct 720  
ggacattaac aaactctatc ttaacattta gatgcaagag cttttatctc actataaatg 780  
cacgatgatt tctcattgtt tctcacaaaa agcattcagt tcattagtcc tacaacaacg 840  
aattcgggtt cccaaatcgc cgccaccatg gcttctatga taccctctc cgctgtgaca 900  
acagtcagcc gtgcctctag ggggcaatcc gccgcagtgg ctccattcgg cggcctcaaa 960  
tccatgactg gattcccagt gaagaaggtc aacttgaca ttacttccat tacaagcaat 1020  
ggtggaagag taaagtgcac gaaaccaact acggttaattg gtgcaggctt cgggtggcctg 1080  
gcactggcaa ttcgtctaca agctgcgggg atccccgtct tactgcttga acaacgtgat 1140

aaacccggcg gtcgggctta tgtctacgag gatcaggggt ttacctttga tgcaggcccc 1200  
acggttatca ccgatcccag tgccattgaa gaactgtttg cactggcagg aaaacagtta 1260  
aaagagtatg tcgaactgct gccggttacg ccgttttacc gcctgtgttg ggagtcaggg 1320  
aaggctcttta attacgataa cgatcaaacc cggtcgaag cgcagattca gcagtttaat 1380  
ccccgcgatg tcgaagggtta tcgtcagttt ctggactatt cacgcgcggt gtttaaagaa 1440  
ggctatctga agctcgggtac tgtccctttt ttatcgttca gagacatgct tcgcgccgca 1500  
cctcaactgg cgaaactgca ggcattggaga agcgtttaca gtaagggtgc cagttacatc 1560  
gaagatgaac atctgcgcca ggcgttttct ttccactcgc tgttggtggg cggaatccc 1620  
ttcgccacct catccattta tacgttgata cacgcgctgg agcgtgagtg gggcgtctgg 1680  
tttccgcgtg gcggcaccgg cgcattagtt caggggatga taaagctgtt tcaggatctg 1740  
ggtggcggaag tcgtgttaaa cgccagagtc agccatatgg aaacgacagg aaacaagatt 1800  
gaagccgtgc atttagagga cggtcgcagg ttcctgacgc aagccgtcgc gtcaaatgca 1860  
gatgtggttc atacctatcg cgacctgtta agccagcacc ctgccgcggt taagcagtcc 1920  
aacaactgc agactaagcg catgagtaac tctctgtttg tgetctattt tggtttgaat 1980  
caccatcatg atcagctcgc gcatcacacg gtttgtttcg gccgcggtta ccgcgagctg 2040  
attgacgaaa tttttaatca tgatggcctc gcagaggact tctcacttta tctgcacgcg 2100  
ccctgtgtca cggattcgtc actgggcgct gaaggttgcg gcagttacta tgtgttggcg 2160  
ccggtgccgc atttaggcac cgcgaacctc gactggacgg ttgagggggc aaaactacgc 2220  
gaccgtatth ttgcgtacct tgagcagcat tacatgcctg gcttacggag tcagctggtc 2280  
acgcaccgga tgtttacgcc gtttgatttt cgcgaccagc ttaatgccta tcatggctca 2340  
gccttttctg tggagcccg tcttaccag agcgcctggt ttcggccgca taaccgcgat 2400  
aaaaccatta ctaatctcta cctggtcggc gcaggcacgc atcccggcg aggcattcct 2460  
ggcgtcatcg gctcggcaaa agcgacagca ggtttgatgc tggaggatct gatttgaggc 2520  
catgcaggcc gatccccgat cgttcaaaca tttggcaata aagtttctta agattgaatc 2580  
ctggtgcggg tcttgcatg attatcatat aatttctggt gaattacgtt aagcatgtaa 2640  
taattaacat gtaatgcatg acgttattta tgagatgggt ttttatgatt agagtccccg 2700  
aattatacat ttaatacgcg atagaaaaca aaatatagcg cgcaaactag gataaattat 2760  
cgcgcgcggt gtcactatg ttactagatc gggccttaat aagcttggtta atcatggtgt 2820  
aggcaaccca aataaaacac caaatatgc acaaggcagt ttgttgatt ctgtagtaca 2880  
gacaaaacta aaagtaatga aagaagatgt ggtgttagaa aaggaaacaa tatcatgagt 2940

-10-

aatgtgtgag cattatggga ccacgaaata aaaagaacat tttgatgagt cgtgtatcct 3000  
cgatgagcct caaaagttct ctcaccccg atagaaacc cttaagcaat gtgcaaagtt 3060  
tgcatctctc actgacataa tgcaaaataa gatatcatcg atgacatagc aactcatgca 3120  
tcatatcatg cctctctcaa cctattcatt cctactcatc tacataagta tcttcagcta 3180  
aatgttagaa cataaaccce taagtcacgt ttgatgagta ttaggcgtga cacatgacaa 3240  
atcacagact caagcaagat aaagcaaat gatgtgtaca taaaactcca gagctatatg 3300  
tcatatgca aaaagaggag agcttataag acaaggcatg actcacaaa attcatttgc 3360  
cttctcgtgc aaaaagagga gggctttaca ttatccatgt catattgcaa aagaaagaga 3420  
gaaagaacaa cacaatgctg cgtcaattat acatatctgt atgtccatca ttattcatcc 3480  
accttctcgtg taccacactt catatatcat gagtcacttc atgtctggac attaacaaac 3540  
tctatcttaa cathtagatg caagagcctt tatctcacta taaatgcacg atgatttctc 3600  
attgtttctc acaaaaagca ttcagttcat tagtcctaca acaacgaatt cggcttccca 3660  
aatcgccgcc accatggcca tcatactcgt acgagcagcg tcgccccggc tctccgccgc 3720  
cgacagcatc agccaccagg ggactctcca gtgctccacc ctgctcaaga cgaagaggcc 3780  
ggcgccgccgc cgggtgatgc cctgctcgtc ccttgacctc caccctggg aggctggccg 3840  
tccctcccc gccgtctact ccagcctcgc cgtcaacccg gcgggagagg ccgtcgtctc 3900  
gtccgagcag aaggctctac acgtcgtgct caagcaggcc gcattgctca aacgccagct 3960  
gcgcacgccg gtctctcagc ccaggcccca ggacatggac atgccacgca acgggctcaa 4020  
ggaagcctac gaccgctcgc gcgagatctg tgaggagtat gccaagacgt ttacctcgg 4080  
aactatgttg atgacagagg agcggcgccg cgccatatgg gccatctatg tgtggtgtag 4140  
gaggacagat gagctttag atgggcaaaa cgccaactac attacacaa cagctttgga 4200  
ccggtgggag aagagacttg aggatctgtt caggggacgt ccttacgaca tgcttgatgc 4260  
cgctctctct gataccatct caaggttccc catagacatt cagccattca gggacatgat 4320  
tgaagggatg aggagtgatc ttaggaagac aagggtataac aacttcgacg agctctacat 4380  
gtactgttac tatgttgctg gaactgtcgg gttaatgagc gtaccagtga tgggcatcgc 4440  
atccgagtct aaagcaacaa ctgaaagcgt gtacagtgtc gccttggtc tcggaattgc 4500  
gaaccaactc acgaacatac tccgggatgt tggagaggat gctagacgag gaaggatata 4560  
tttaccacaa gatgagcttg cacaggcagg gctctctgat gaggacatct tcaaaggggt 4620  
cgtcacgaac cgggtggagaa acttcatgaa gaggcagatc aagagggcca ggatgttttt 4680  
tgaggaggca gagagagggg taactgagct ctcacaggct agcagatggc cagtatgggc 4740

-11-

ttccctggtg ttgtacaggc agatcctgga tgagatcgaa gccaacgact acaacaactt 4800  
 cacgaagagg gcgtatggtg gtaaagggaa gaagttgcta gcacttctg tggcatatgg 4860  
 aaaatcgcta ctgctcccat gttcattgag aaatggccag acctagggcc atgcaggccg 4920  
 atccccgatc gttcaaacat ttggcaataa agtttcttaa gattgaatcc tgttgccggt 4980  
 cttgcatga ttatcatata atttctggtg aattacgtta agcatgtaat aattaacatg 5040  
 taatgcatga cgttatztat gagatgggtt tttatgatta gagtcccgca attatacatt 5100  
 taatacgca tagaaaaaa aatatagcgc gcaaactagg ataaattatc gcgcgcggtg 5160  
 tcatttatgt tactagatcg 5180

&lt;210&gt; 4

&lt;211&gt; 5180

&lt;212&gt; DNA

&lt;213&gt; SYNTHETIC - 12424

&lt;400&gt; 4

gttaatcatg gtgtaggcaa cccaaataaa acacaaaaat atgcacaagg cagtttggtg 60  
 tattctgtag tacagacaaa actaaaagta atgaaagaag atgtggtggt agaaaaggaa 120  
 acaatatcat gagtaatgtg tgagcattat gggaccacga aataaaaaga acattttgat 180  
 gagtctgtga tcctcgatga gcctcaaaag ttctctcacc cgggataaga aacccttaag 240  
 caatgtgcaa agtttgcatt ctccactgac ataatgcaaa ataagatatc atcgatgaca 300  
 tagcaactca tgcattcatat catgcctctc tcaacctatt cattcctact catctacata 360  
 agtatcttca gctaaatggt agaacataaa ccataagtc acgtttgatg agtattaggc 420  
 gtgacacatg acaaatcaca gactcaagca agataaagca aaatgatgtg tacataaaac 480  
 tccagagcta tatgtcatat tgcaaaaaga ggagagctta taagacaagg catgactcac 540  
 aaaaattcat ttgcctttcg tgtcaaaaag aggagggtt tacattatcc atgtcatatt 600  
 gcaaaagaaa gagagaaaga acaacacaat gctgcgtcaa ttatacatat ctgtatgtcc 660  
 atcattatcc atccaccttt cgtgtaccac acttcatata tcatgagtca cttcatgtct 720  
 ggacattaac aaactctatc ttaacattta gatgcaagag cttttatctc actataaatg 780  
 cactgatgatt tctcattggt tctcacaana agcattcagt tcattagtcc tacaacaacg 840  
 aattcgggtt cccaaatcgc cgccaccatg gcttctatga tatcctcttc cgctgtgaca 900  
 acagtcagcc gtgcctctag ggggcaatcc gccgcagtgg ctccattcgg cggcctcaaa 960

-12-

tccatgactg gattcccagt gaagaaggtc aacactgaca ttacttccat tacaagcaat 1020  
gggtggaagag taaagtgcac gaaaccaact acggtaattg gtgcaggcctt cgggtggcctg 1080  
gcactggcaa ttcgctctaca agctgcgggg atccccgtct tactgcttga acaacgtgat 1140  
aaacccggcg gtccggctta tgtctacgag gatcaggggt ttacctttga tgcaggcccg 1200  
acggttatca ccgatcccag tgccattgaa gaactgtttg cactggcagg aaaacagtta 1260  
aaagagtatg tcgaactgct gccggttacg ccgttttacc gcctgtgttg ggagtcaggg 1320  
aaggtcttta attacgataa cgatcaaacc cggctcgaag cgcagattca gcagtttaat 1380  
ccccgcgatg tcgaaggcta tcgtcagttt ctggactatt cacgcgcggt gtttaaagaa 1440  
ggctatctga agctcggtag tgtccctttt ttatcgttca gagacatgct tcgcgccgca 1500  
cctcaactgg cgaaactgca ggcatggaga agcgtttaca gtaagggtgc cagttacatc 1560  
gaagatgaac atctgcgcca ggcgttttct ttccactcgc tgttggtggg cggcaatccc 1620  
ttcgccacct catccattta tacgttgata cacgcgctgg agcgtgagtg gggcgtcttg 1680  
tttcgcgctg gcggcaccgg cgcattagtt caggggatga taaagctgtt tcaggatctg 1740  
gggtggcgaag tcgtgttaaa cgccagagtc agccatatgg aaacgacagg aaacaagatt 1800  
gaagccgtgc atttagagga cggtcgcagg ttccctgacgc aagccgtcgc gtcaaagtca 1860  
gatgtggttc atacctatcg cgacctgtta agccagcacc ctgccgcggt taagcagtcc 1920  
aacaaaactgc agactaagcg catgagtaac tctctgtttg tgctctatct tggtttgaat 1980  
caccatcatg atcagctcgc gcatcacacg gtttgtttcg gccgcggtta ccgcgagctg 2040  
attgacgaaa tttttaatca tgatggcctc gcagaggact tctcacttta tctgcacgcg 2100  
ccctgtgtca cggattcgtc actggcgctt gaagggtgcg gcagttacta tgtgttggcg 2160  
ccggtgcgcg atttaggcac cgcgaacctc gactggacgg ttgaggggccc aaaactacgc 2220  
gaccgtatct ttgcgtacct tgagcagcat tacatgcttg gcttacggag tcagctggtc 2280  
acgcaccgga tgtttacgcc gtttgatctt cgcgaccagc ttaatgccta tcatggctca 2340  
gccttttctg tggagcccggt tcttaccag agcgcctggt ttcggccgca taaccgcgat 2400  
aaaaccatta ctaatctcta cctggtcggc gcaggcacgc atccccggcg aggcattcct 2460  
ggcgtcatcg gctcggcaaa agcgacagca ggtttgatgc tggaggatct gatttgaggc 2520  
catgcaggcc gatccccgat cgttcaaaca tttggcaata aagtttctta agattgaatc 2580  
ctgttgccgg tcttgcatg attatcatat aatttctgtt gaattacgtt aagcatgtaa 2640  
taattaacat gtaatgcag acgttatctta tgagatgggt ttttatgatt agagtcccg 2700  
aattatacat ttaatacgcg atagaaaaca aaatatagcg cgcaaactag gataaattat 2760



-13-

cgcgcgcggt gtcacatctatg ttactagatc gggccttaat aagcttggtta atcatgggtgt	2820
aggcaaccca aataaaacac caaaatatgc acaaggcagt ttgttgatt ctgtagtaca	2880
gacaaaacta aaagtaatga aagaagatgt ggtgtagaa aaggaaacaa tatcatgagt	2940
aatgtgtgag cattatggga ccacgaaata aaaagaacat tttgatgagt cgtgtatcct	3000
cgatgagcct caaaagttct ctcaccccg ataagaaacc cttaagcaat gtgcaaagtt	3060
tgcattctcc actgacataa tgcaaaataa gatatcatcg atgacatagc aactcatgca	3120
tcatatcatg cctctctcaa cctattcatt cctactcatc tacataagta tcttcagcta	3180
aatgttagaa cataaaccca taagtcacgt ttgatgagta ttaggcgtga cacatgacaa	3240
atcacagact caagcaagat aaagcaaaat gatgtgtaca taaaactcca gagctatatg	3300
tcatatgca aaaagaggag agcttataag acaaggcatg actcacaaaa attcatttgc	3360
ctttcgtgtc aaaaagagga gggctttaca ttatccatgt catattgcaa aagaagaga	3420
gaaagaacaa cacaatgctg cgtcaattat acatatctgt atgtccatca ttattcatcc	3480
acctttcgtg taccacactt catatatcat gagtcacttc atgtctggac attaacaac	3540
tctatcttaa catttagatg caagagcctt tatctcacta taaatgcacg atgatttctc	3600
attgtttctc acaaaaagca ttcagttcat tagtcctaca acaacgaatt cggcttccca	3660
aatcgccgcc accatggcca tcatactgt acgagcagcg tcgccggggc tctccgccgc	3720
cgacagcatc agccaccagg ggactctcca gtgctccacc ctgctcaaga cgaagaggcc	3780
ggcgccgcgg cggtggatgc cctgctcgct ccttggcctc caccgctggg aggctggccg	3840
tccctcccc gccgtctact ccagcctgcc cgtcaaccgc gcgggagagg ccgtcgtctc	3900
gtccgagcag aaggtctacg acgtcgtgct caagcaggcc gcattgctca aacgccagct	3960
gcgcacgccg gtcctcgacg ccaggcccca ggacatggac atgccacgca acgggctcaa	4020
ggaagcctac gaccgctgcg gcgagatctg tgaggagtat gccaaagcgt tttacctcg	4080
aactatgttg atgacagagg agcggcgccg cgccatatgg gccatctatg tgtggtgtag	4140
gaggacagat gagcttgtag atgggccaaa cgccaactac attacaccaa cagctttgga	4200
ccggtgggag aagagacttg aggatctggt cacgggacgt ccttacgaca tgcttgatgc	4260
cgctctctct gataccatct caaggttccc catagacatt cagccattca gggacatgat	4320
tgaagggatg aggagtgatc ttaggaagac aaggtataac aacttcgacg agctctacat	4380
gtactgctac tatgttgctg gaactgtcgg gttaatgagc gtacctgtga tgggcatcgc	4440
aaccgagtct aaagcaacaa ctgaaagcgt atacagtgc gccttggtc tggaattgc	4500
gaaccaactc acgaacatac tccgggatgt tggagaggat gctagaagag gaaggatata	4560

-14-

tttaccacaa gatgagcttg cacaggcagg gctctctgat gaggacatct tcaaaggggt 4620  
cgtcacgaac cggtaggagaa acttcatgaa gaggcagatc aagagggcca ggatgttttt 4680  
tgaggaggca gagagagggg taactgagct ctcacaggct agcagatggc cagtatgggc 4740  
ttccctgttg ttgtacaggc agatcctgga tgagatcgaa gccaacgact acaacaactt 4800  
cacgaagagg gcgtatgttg gtaaagggaa gaagttgcta gcacttcctg tggcatatgg 4860  
aaaatcgcta ctgctcccat gttcattgag aaatggccag acctagggcc atgcaggccg 4920  
atccccgatc gttcaaacat ttggcaataa agtttcttaa gattgaatcc tgttgccggg 4980  
cttgcatga ttatcatata atttctgttg aattacgtta agcatgtaat aattaacatg 5040  
taatgcatga cgttatttat gagatgggtt tttatgatta gagtcccgca attatacatt 5100  
taatacgca tagaaaacaa aatatagcgc gcaaactagg ataaattatc gcgcgcgggtg 5160  
tcatttatgt tactagatcg 5180

&lt;210&gt; 5

&lt;211&gt; 5653

&lt;212&gt; DNA

&lt;213&gt; SYNTHETIC

<400> 5  
gttaatcatg gtgtaggcaa cccaaataaa acaccaaagt atgcacaagg cagtttggtg 60  
tattctgtag tacagacaaa actaaaagta atgaaagaag atgtgggtgt agaaaaggaa 120  
acaatatcat gagtaatgtg tgagcattat gggaccacga aataaaaaga acattttgat 180  
gagtcgtgta tcctcgatga gcctcaaaag ttctctcacc ccggataaga aacccttaag 240  
caatgtgcaa agtttgcatc ctccactgac ataatgcaaa ataagatatc atcgatgaca 300  
tagcaactca tgcacatat catgcctctc tcaacctatt cattcctact catctacata 360  
agtatcttca gctaaatgtt agaacataaa ccataagtc acgtttgatg agtattaggc 420  
gtgacacatg acaaatcaca gactcaagca agataaagca aaatgatgtg tacataaaac 480  
tccagagcta tatgtcatat tgcaaaaaga ggagagctta taagacaagg catgactcac 540  
aaaaattcat ttgcctttcg tgtcaaaaag aggagggcct tacattatcc atgtcatatt 600  
gcaaaagaaa gagagaaaga acaacacaat gctgcgtaaa ttatacatat ctgtatgtcc 660  
atcattatcc atccaccttt cgtgtaccac acttcatata tcatgagtca cttcatgtct 720  
ggacattaac aaactctatc ttaacattta gatgcaagag cttttatctc actataaatg 780

-15-

cacgatgatt tctcattggt tctcacaaaa agcattcagt tcattagtcc tacaacaacg 840  
aattcgggctt cccgggtaca gggtaaattt ctagtttttc tccttcattt tcttggttag 900  
gacccttttc tctttttatt tttttgagct ttgatctttc tttaaactga tctatttttt 960  
aattgattgg ttatcgtgta aatattacat agctttaact gataatctga ttactttatt 1020  
tcgtgtgtct ttgatcatct tgatagttac agaaccgtcg actctagaga agccatttaa 1080  
atcgccgcca ccatggcttc tatgatatcc tcttcgctg tgacaacagt cagccgtgcc 1140  
tctagggggc aatccgccgc agtgggtcca ttcggcggcc tcaaatecat gactggattc 1200  
ccagtgaaga aggtcaacac tgacattact tccattacaa gcaatgggtg aagagtaaag 1260  
tgcatggcgg ccgccaacc aactacggta attggtgcag gcttcgggtg cctggcactg 1320  
gcaattcgtc tacaagctgc ggggatcccc gtcttactgc ttgaacaacg tgataaaccc 1380  
ggcggtcggg cttatgtcta cgaggatcag ggggtttacct ttgatgcagg cccgacgggt 1440  
atcaccgatc ccagtgccat tgaagaactg tttgcactgg caggaaaaca gttaaaagag 1500  
tatgtcgaac tgctgccggt tacgccgttt taccgcctgt gttgggagtc aggggaaggtc 1560  
tttaattacg ataacgatca aaccggctc gaagcgcaga ttcagcagtt taatccccgc 1620  
gatgtcgaag gttatcgtca gtttctggac tattcacgcg cgggtgttaa agaaggctat 1680  
ctgaagctcg gtactgtccc ttttttatcg ttcagagaca tgcttcgcgc cgcacctcaa 1740  
ctggcgaaac tgcaggcatg gagaagcgtt tacagtaagg ttgccagtta catcgaagat 1800  
gaacatctgc gccaggcgtt ttctttccac tcgctgttg tgggcggcaa tcccttcgcc 1860  
acctcatcca tttatacgtt gatacacgcg ctggagcgtg agtggggcgt ctggtttccg 1920  
cgtggcggca ccggcgcatt agttcagggg atgataaagc tgtttcagga tctgggtggc 1980  
gaagtcgtgt taaacgccag agtcagccat atggaaacga caggaaacaa gattgaagcc 2040  
gtgcatttag aggacggctg caggttcctg acgcaagccg tcgcgtcaaa tgcagatgtg 2100  
gttcatacct atcgcgacct gttaagccag caccctgccg cggttaagca gtccaacaaa 2160  
ctgcagacta agcgcgatg taactctctg tttgtgtctt attttggttt gaatcaccat 2220  
catgatcagc tcgcgcatca cacggtttgt ttcggcccgc gttaccgcga gctgattgac 2280  
gaaattttta atcatgatgg cctcgcagag gacttctcac tttatctgca cgcgccctgt 2340  
gtcacggatt cgtcactggc gcctgaaggt tgcggcagtt actatgtgtt ggcgccgggtg 2400  
ccgcatttag gcaaccgcga cctcgactgg acggttgagg ggccaaaact acgcgaccgt 2460  
atttttgcgt accttgagca gcattacatg cctggcttac ggagtcagct ggtcacgcac 2520  
cggatgttta cgcggtttga ttttcgcgac cagcttaatg cctatcatgg ctcagccttt 2580

-16-

tctgtggagc ccgttcttac ccagagcgcc tggtttcggc cgcataaccg cgataaaacc 2640  
attactaadc tctacctggt cggcgcaggc acgcatcccg gcgcaggcat tcctggcgctc 2700  
atcggctcgg caaaagcgac agcaggtttg atgctggagg atctgatttg aggtacctcg 2760  
acggccatgc aggcgatcc ccgatcgttc aaacatttgg caataaagtt tcttaagatt 2820  
gaatcctggt gccggctctg cgatgattat catataattt ctgttgaatt acgttaagca 2880  
tgtaataatt aacatgtaat gcatgacgtt atttatgaga tgggttttta tgattagagt 2940  
cccgaatta tacatttaac acgcgataga aaacaaaata tagcgcgcaa actaggataa 3000  
attatcgcg cgggtgtcat ctatgttact agatcgggcc ttaatcgcaa gcttgtaatt 3060  
catggtgtag gcaacccaaa taaaacacca aaatatgcac aaggcagttt gttgtattct 3120  
gtagtacaga caaaactaaa agtaatgaaa gaagatgtgg tgtagaaaa ggaaacaata 3180  
tcatgagtaa tgtgtgagca ttatgggacc acgaaataaa aagaacattt tgatgagtcg 3240  
tgtatcctcg atgagcctca aaagttctct cccccggat aagaaccct taagcaatgt 3300  
gcaaagtttg cattctccac tgacataatg caaaataaga tatcatcgat gacatagcaa 3360  
ctcatgcac atatcatgcc tctctcaacc tattcattcc tactcatcta cataagtac 3420  
ttcagctaaa tgtagaaca taaaccata agtcacgtt gatgagtatt aggcgtgaca 3480  
catgacaaat cacagactca agcaagataa agcaaatga tgtgtacata aaactccaga 3540  
gctatatgtc atattgcaaa aagaggagag cttataagac aaggcatgac tcacaaaaat 3600  
tcatttgcct ttcgtgtcaa aaagaggagg gctttacatt atccatgtca tattgcaaaa 3660  
gaaagagaga aagaacaaca caatgctgcg tcaattatac atatctgtat gtccatcatt 3720  
attcatccac ctttcgtgta ccacacttca tatatcatga gtcactcat gtctggacat 3780  
taacaaactc tatcttaaca tttagatgca agagcctta tctactata aatgcacgat 3840  
gatttctcat tgtttctcac aaaaagcatt cagttcatta gtctacaac aacgaattcg 3900  
gcttcccggt tacagggtaa atttctagtt tttctccttc attttcttgg ttaggacct 3960  
tttctctttt tatttttttg agctttgatc tttctttaa ctgatctatt ttttaattga 4020  
ttggttatcg tgtaaatatt acatagcttt aactgataat ctgattactt tatttctgt 4080  
gtctttgatc atcttgatag ttacagaacc gtcgactcta gagaagccat ttaaactgcc 4140  
gccaccatgg ccatacact cgtacgagca gcgtcgccgg ggctctccgc cgccgacagc 4200  
atcagccacc aggggactct ccagtgtcc accctgtca agacgaagag gccggcggcg 4260  
cggcggtgga tgccctgctc gtccttggc ctccaccgt gggaggctgg ccgtccctcc 4320  
cccgcgtct actccagcct gcccgtaac ccggcgggag aggcgctgt ctgctccgag 4380

-17-

cagaaggtct acgacgtcgt gctcaagcag gccgcattgc tcaaacgccca gctgcgcacg 4440  
 ccggtcctcg acgccaggcc ccaggacatg gacatgccac gcaacgggct caaggaagcc 4500  
 tacgaccgct gcggcgagat ctgtgaggag tatgccaaaga cgttttacct cggaactatg 4560  
 ttgatgacag aggagcggcg ccgcgccata tgggccatct atgtgtggtg taggaggaca 4620  
 gatgagcttg tagatgggcc aaacgccaac tacattacac caacagcttt ggaccgggtg 4680  
 gagaagagac ttgaggatct gttcacggga cgtccttacg acatgcttga tgccgctctc 4740  
 tctgatacca tctcaagggt ccccatagac attcagccat tcaggggacat gattgaaggg 4800  
 atgaggagtg atcttaggaa gacaagggtat aacaacttcg acgagctcta catgtactgc 4860  
 tactatgttg ctggaactgt cgggttaatg agcgtacctg tgatgggcat cgcaaccgag 4920  
 tctaaagcaa caactgaaag cgtatacagt gctgccttgg ctctgggaat tgccaaccaa 4980  
 ctcacgaaca tactccggga tggttgagag gatgctagaa gaggaaggat atatttacca 5040  
 caagatgagc ttgcacaggc agggctctct gatgaggaca tcttcaaagg ggtcgtcacg 5100  
 aaecgggtgga gaaacttcat gaagaggcag atcaagaggg ccaggatgtt ttttgaggag 5160  
 gcagagagag gggtaaata gctctcacag gctagcagat ggccagtatg ggcttccttg 5220  
 ttgttgtaca ggcagatcct ggatgagatc gaagccaacg actacaacaa cttcacgaag 5280  
 agggcgatg ttggtaaagg gaagaagttg ctagcacttc ctgtggcata tggaaaatcg 5340  
 ctactgctcc catgttcatt gagaaatggc cagacctagg gccatgcagg ccgatccccg 5400  
 atcgttcaaa catttgga taaagtttct taagattgaa tctgtttgcc ggtcttgcca 5460  
 tgattatcat ataatttctg ttgaattacg ttaagcatgt aataattaac atgtaatgca 5520  
 tgacgttatt tatgagatgg gtttttatga ttagagtccc gcaattatac atttaatacg 5580  
 cgatagaaaa caaatatag cgcgcaaact aggataaatt atcgcgcgcg gtgtcatcta 5640  
 tgttactaga tcg 5653

&lt;210&gt; 6

&lt;211&gt; 5714

&lt;212&gt; DNA

&lt;213&gt; SYNTHETIC - 11586

&lt;400&gt; 6

gttaatcatg gtgtaggcaa cccaaataaa acaccaaagt atgcacaagg cagtttggtg 60  
 tattctgtag tacagacaaa actaaaagta atgaaagaag atgtggtggt agaaaaggaa 120

-18-

acaatatcat gagtaatgtg tgagcattat gggaccacga aataaaaaga acattttgat 180  
 gagtcgtgta tcctcgatga gcctcaaaaag ttctctcacc ccggataaga aacccttaag 240  
 caatgtgcaa agtttgcatc ctccactgac ataatgcaaa ataagatata atcgatgaca 300  
 tagcaactca tgcacatata catgcctctc tcaacctatt cattcctact .catctacata 360  
 agtatcttca gctaaatgtt agaacataaa cccataagtc acgtttgatg agtattaggc 420  
 gtgacacatg acaaatcaca gactcaagca agataaagca aaatgatgtg tacataaaac 480  
 tccagagcta tatgtcatat tgcaaaaaga ggagagctta taagacaagg catgactcac 540  
 aaaaattcat ttgcctttcg tgtcaaaaag aggagggtt tacattatcc atgtcatatt 600  
 gcaaaagaaa gagagaaaga acaacacaat gctgctcaa ttatacatat ctgtatgtcc 660  
 atcattatcc atccaccttt cgtgtaccac acttcatata tcatgagtca ctctatgtct 720  
 ggacattaac aaactctatc ttaacattta gatgcaagag cctttatctc actataaatg 780  
 cacgatgatt tctcattgtt tctcacaaaa agcattcagt tcattagtcc tacaacaacg 840  
 aattcgggtt cccgggtaca gggtaaattt ctagtttttc tccttcattt tcttgggttag 900  
 gacccttttc tctttttatt tttttgagct ttgatctttc tttaaactga tctatttttt 960  
 aattgattgg ttatcgtgta aatattacat agctttaact gataatctga ttactttatt 1020  
 tcgtgtgtct ttgatcatct tgatagttac agaaccgtcg actctagaga agccatttaa 1080  
 atcgccgcca ccatggcttc tatgatatac tcttcgctg tgacaacagt cagccgtgcc 1140  
 tctagggggc aatccgccgc agtgggtcca ttcggcgccc tcaaatccat gactggattc 1200  
 ccagtgaaga aggtcaacac tgacattact tccattacaa gcaatgggtg aagagtaaag 1260  
 tgcatggcgg ccgccaacc aactacggta attgggtgcag gcttcgggtg cctggcactg 1320  
 gcaattcgtc tacaagctgc ggggatcccc gtcttactgc ttgaacaacg tgataaacc 1380  
 ggcggctcggg cttatgtcta cgaggatcag gggtttacct ttgatgcagg cccgacggtt 1440  
 atcaccgatc ccagtgccat tgaagaactg tttgactgg caggaaaaca gttaaaagag 1500  
 tatgtgcaac tgctgccgtt tacgccgttt taccgcctgt gttgggagtc aggaaggtc 1560  
 ttaattacg ataacgatca aaccggctc gaagcgaga ttcagcagtt taatccccgc 1620  
 gatgtcgaag gttatcgtca gtttctggac tattcacgcg cggtgtttta agaaggctat 1680  
 ctgaagctcg gtactgtccc ttttttatcg ttcagagaca tgcttcgcgc cgcacctcaa 1740  
 ctggcgaaac tgcaggcatg gagaagcgtt tacagtaagg ttgccagtt catcgaagat 1800  
 gaacatctgc gccaggcgtt ttctttccac tcgctgttg tggcgcgcaa tcccttcgcc 1860  
 acctcatcca ttatacgtt gatacacgcg ctggagcgtg agtggggcgt ctggtttccg 1920

-19-

cgtggcggca ccggcgcat agttcagggg atgataaagc tgtttcagga tctgggtggc 1980  
gaagtcgtgt taaacgccag agtcagccat atggaaacga caggaaacaa gattgaagcc 2040  
gtgcatttag aggacggtcg caggttcctg acgcaagccg tcgctcaaa tgcagatgtg 2100  
gttcatacct atcgcgacct gttaagccag caccctgccg cggttaagca gtccaacaaa 2160  
ctgcagacta agcgcatgag taactctctg tttgtgtctt attttggttt gaatcaccat 2220  
catgatcagc tcgcgcatca cacggtttgt ttcggcccg gttaccgca gctgattgac 2280  
gaaatTTTTA atcatgatgg cctcgagag gacttctcac tttatctgca cgcgcctgt 2340  
gtcacggatt cgtcactggc gcctgaaggt tgcggcagtt actatgtgtt ggcgccgggtg 2400  
ccgcatttag gcaccgcgaa cctcgactgg acggttgagg ggccaaaact acgcgaccgt 2460  
atTTTTgcgt accttgagca gcattacatg cctggcttac ggagtcagct ggtcacgcac 2520  
cggatgttta cgccgtttga ttttcgcgac cagcttaatg cctatcatgg ctcagccttt 2580  
tctgtggagc ccgttcttac ccagagcgcc tggtttcggc cgcataaccg cgataaaacc 2640  
attactaatc tctacctggg cggcgagggc acgcatcccg gcgcaggcat tcctggcgtc 2700  
atcggtcgg caaaagcgac agcaggtttg atgctggagg atctgatttg aggtacctcg 2760  
acggccatgc aggccgatcc ccgatcgctc aaacatttg caataaagtt tcttaagatt 2820  
gaatcctgtt gccggctctg cgatgattat catataattt ctgttgaatt acgttaagca 2880  
tgtaataatt aacatgtaat gcatgacgtt atttatgaga tgggttttta tgattagagt 2940  
cccgcatta taatTTAat acgcgataga aaacaaaata tagcgcgcaa actaggataa 3000  
attatcgcgc gcggtgtcat ctatgttact agatcggggc ttaaaactga aggcgggaaa 3060  
cgacaatctg atctctagga agcttggtta tcatggtgta ggcaacccaa ataaaacacc 3120  
aaaatatgca caaggcagtt tgttgatttc tgtagtacag acaaaactaa aagtaatgaa 3180  
agaagatgtg gtgttagaaa aggaaacaat atcatgagta atgtgtgagc attatgggac 3240  
cacgaaataa aaagaacatt ttgatgagtc gtgtatcttc gatgagcctc aaaagttctc 3300  
tcaccccgga taagaaaccc ttaagcaatg tgcaaagttt gcattctcca ctgacataat 3360  
gcaaaataag atatcatcga tgacatagca actcatgcat catatcatgc ctctctcaac 3420  
ctattcatte ctactcatct acataagtat cttcagctaa atgttagaac ataaacccat 3480  
aagtcacgtt tgatgagtat taggcgtgac acatgacaaa tcacagactc aagcaagata 3540  
aagcaaaatg atgtgtacat aaaactccag agctatatgt catattgcaa aaagaggaga 3600  
gcttataaga caaggcatga ctcacaaaaa ttcatttgcc tttcgtgtca aaaagaggag 3660  
ggctttacat tatccatgtc atattgcaaa agaaagagag aaagaacaac acaatgctgc 3720

gtcaattata catatctgta tgtccatcat tattcatcca cctttcgtgt accacacttc 3780  
atatatcatg agtcacttca tgtctggaca ttaacaaact ctatcttaac atttagatgc 3840  
aagagccttt atctcactat aaatgcacga tgatttctca ttgtttctca caaaaagcat 3900  
tcagttcatt agtcctacaa caacgaattc ggcttcccggtg gtacagggtg aatttctagt 3960  
ttttctcctt cattttcttg gttaggaccc ttttctcttt ttattttttt gagctttgat 4020  
ctttctttta actgatctat tttttaattg attggttctc gtgtaaatat tacatagctt 4080  
taactgataa tctgattact ttatttcgtg tgtctttgat catcttgata gttacagaac 4140  
cgctcgactct agagaageca tttaaatcgc cgccaccatg gcggccatca cgctcctacg 4200  
ttcagcgtct cttccggggc tctccgacgc cctcgcccggtg gacgctgctg ccgtccaaca 4260  
tgtctgctcc tcctacctgc ccaacaacaa ggagaagaag aggaggtgga tcctctgctc 4320  
gctcaagtac gcctgccttg gcgtcgaccc tgccccgggc gagattgccc ggacctcgcc 4380  
gggtgtactcc agcctcaccg tcaccctgc tggagaggcc gtcctctcct cggagcagaa 4440  
gggtgtacgac gtcgtcctca agcaggcagc attgctcaaa cgccacctgc gccacaacc 4500  
acacaccatt cccatcgctc ccaaggacct ggacctgcca agaaacggcc tcaagcaggc 4560  
ctatcatcgc tgcggagaga tctgcgagga gtatgccaa agccttttacc ttggaactat 4620  
gctcatgacg gaggaccgac ggcgcgccat atgggccatc tatgtgtggt gtaggaggac 4680  
agatgagctt gtagatggac caaatgcctc gcacatcaca ccgtcagccc tggaccggtg 4740  
ggagaagagg cttgatgatc tcttcaccgg acgcccctac gacatgcttg atgctgact 4800  
ttctgatacc atctccaagt ttctataga tattcagcct ttcagggaca tgatagaagg 4860  
gatgcgggtca gacctcagaa agactagata caagaacttc gacgagctct acatgtactg 4920  
ctactatggt gctggaactg tggggctaata gagggttctc gtgatgggtg ttgcaccga 4980  
gtcgaaggca acaactgaaa gtgtgtacag tgctgctttg gctctcgga ttgcaaacca 5040  
gctcacaat atactccgtg acgttgagga ggacgcgaga agaggaggga tatatttacc 5100  
acaagatgaa cttgcagagg cagggctctc tgatgaggac atcttcaatg gcgttgtgac 5160  
taacaaatgg agaagcttca tgaagagaca gatcaagaga gctaggatgt tttttgagga 5220  
ggcagagaga ggggtgaccg agctcagcca ggcaagccgg tggccggtct gggcgtctct 5280  
gttggttatac cggcaaatcc ttgacgagat agaagcaaac gattacaaca acttcacaaa 5340  
gagggcgtag gttgggaagg cgaagaaatt gctagcgctt ccagttgcat atggtagatc 5400  
attgctgatg cctactcac tgagaaatag ccagaagtag ggccatgcag gccgatcccc 5460  
gatcgttcaa acatttgga ataaagtctc ttaagattga atcctgttgc cggcttgcg 5520



-21-

atgattatca tataatcttct gttgaattac gtttaagcatg taataattaa catgtaatgc 5580  
 atgacgttat ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac 5640  
 gcgatagaaa acaaaatata gcgcgcaaac taggataaat tatcgcgcgcg ggtgtcatct 5700  
 atgttactag atcg 5714

&lt;210&gt; 7

&lt;211&gt; 5974

&lt;212&gt; DNA

&lt;213&gt; SYNTHETIC - 7651

&lt;400&gt; 7

gttaatcatg gtgtaggcaa cccaaataaa acaccaaaat atgcacaagg cagtttggtg 60  
 tattctgtag tacagacaaa actaaaagta atgaaagaag atgtggtgtt agaaaaggaa 120  
 acaatatcat gagtaatgtg tgagcattat gggaccacga aataaaaaga acattttgat 180  
 gagtcgtgta tcctcgatga gcctcaaaag ttctctcacc ccggataaga aacccttaag 240  
 caatgtgcaa agtttgcatt ctccactgac ataatgcaaa ataagatata atcgatgaca 300  
 tagcaactca tgcatacat catgcctctc tcaacctatt cattcctact catctacata 360  
 agtatcttca gctaaatgtt agaacataaa cccataagtc acgtttgatg agtattaggc 420  
 gtgacacatg acaaatcaca gactcaagca agataaagca aaatgatgtg tacataaaac 480  
 tccagagcta tatgtcatat tgcaaaaaga ggagagctta taagacaagg catgactcac 540  
 aaaaattcat ttgcctttcg tgtcaaaaag aggagggtt tacattatcc atgtcatatt 600  
 gcaaaaagaaa gagagaaaga acaacacaat gctgcgtcaa ttatacatat ctgtatgtcc 660  
 atcattattc atccaccttt cgtgtaccac acttcatata tcatgagtca cttcatgtct 720  
 ggacattaac aaactctatc ttaacattta gatgcaagag cttttatctc actataaatg 780  
 cacgatgatt tctcattgtt tctcacaaa agcattcagt tcattagtcc tacaacaacg 840  
 aattcggtt cccgggtaca gggtaaattt ctagtttttc tccttcattt tcttggttag 900  
 gacccttttc tctttttatt tttttgagct ttgatctttc tttaaactga tctatttttt 960  
 aattgattgg ttatcgtgta aatattacat agctttaact gataatctga ttactttatt 1020  
 tcgtgtgtct ttgatcatct tgatagttac agaaccgtcg actctagaga agccatttaa 1080  
 atcgccgcca ccatggcttc tatgatatac tcttcgctg tgacaacagt cagccgtgcc 1140  
 tctagggggc aatccgccc agtgggtcca ttcggcggcc tcaaatccat gactggattc 1200

-22-

ccagtgaaga aggtcaacac tgacattact tccattacaa gcaatggtgg aagagtaaag 1260  
tgcatggcgg cgcgcaaacc aactacggta attggtgcag gcttcggtgg cctggcactg 1320  
gcaattcgtc tacaagctgc ggggatcccc gtcttactgc ttgaacaacg tgataaacc 1380  
ggcggtcggg cttatgtcta cgaggatcag ggggtttacct ttgatgcagg cccgacgggt 1440  
atcacccgatc ccagtgccat tgaagaactg tttgcactgg caggaaaaca gttaaaagag 1500  
tatgtcgaac tgctgccggg tacgccgttt taccgcctgt gttgggagtc agggaaggtc 1560  
tttaattacg ataacgatca aaccggctc gaagcgaga ttcagcagtt taatccccgc 1620  
gatgtcgaag gttatcgta gtttctggac tattcacgcg cgggtgttta agaaggctat 1680  
ctgaagctcg gtactgtccc ttttttatcg ttcagagaca tgcttcgcgc cgcacctcaa 1740  
ctggcgaaac tgcaggcatg gagaagcgtt tacagtaagg ttgccagtta catcgaagat 1800  
gaacatctgc gccaggcgtt ttctttccac tcgctgttgg tgggcggcaa tcccttcgcc 1860  
acctcatcca tttatacgtt gatacacgcg ctggagcgtg agtggggcgt ctggtttccg 1920  
cgtggcggca ccggcgcat agttcagggg atgataaagc tgtttcagga tctgggtggc 1980  
gaagtcgtgt taaacgccag agtcagccat atggaaacga caggaaacaa gattgaagcc 2040  
gtgcatttag aggacggtcg cagggttcctg acgcaagccg tcgctcaaa tgcagatgtg 2100  
gttcatacct atcgcgacct gttaagccag caccctgccg cggttaagca gtccaacaaa 2160  
ctgcagacta agcgcatgag taactctctg tttgtgctct attttggtt gaatcaccat 2220  
catgatcagc tcgcgcata cagggtttgt ttccggccgc gttaccgca gctgattgac 2280  
gaaattttta atcatgatgg cctcgagag gacttctcac tttatctgca cgcgccctgt 2340  
gtcacggatt cgtcactggc gcctgaaggt tgcggcagtt actatgtgtt ggcgccggtg 2400  
ccgcatttag gcaccgcgaa cctcgactgg acggttgagg ggccaaaact acgcgaccgt 2460  
atthttgcgt accttgagca gcattacatg cctggcttac ggagtcagct ggtcacgcac 2520  
cggatgttta cgcggtttga ttttcgcgac cagcttaatg cctatcatgg ctcagccttt 2580  
tctgtggagc ccgttcttac ccagagcgcc tggtttcggc cgcataaccg cgataaaacc 2640  
attactaatc tctacctggt cggcgcaggc acgcatcccc gcgcaggcat tcctggcgtc 2700  
atcggctcgg caaaagcgac agcaggtttg atgctggagg atctgatttg aggtacctcg 2760  
acggccatgc aggcgatcc ccgatcggtc aaacatttgg caataaagtt tcttaagatt 2820  
gaatcctgtt gccggtcttg cgatgattat catataattt ctgttgaatt acgttaagca 2880  
tgtaataatt aacatgtaat gcatgacgtt atttatgaga tgggttttta tgattagagt 2940  
cccgaatta tacatttaac acgcgataga aaacaaaata tagcgcgcaa actaggataa 3000

-23-

attatcgcg gcggtgtcat ctatgttact agatcgggcc ttaatgttcg gggcgaacat 3060  
 cgcaagcttg ttaatcatgg tgtaggcaac ccaaataaaa caccaaaata tgcacaaggg 3120  
 agtttgttgt attctgtagt acagacaaaa ctaaaagtaa tgaaagaaga tgtggtgtta 3180  
 gaaaaggaaa caatatcatg agtaatgtgt gagcattatg ggaccacgaa ataaaaagaa 3240  
 cattttgatg agtcgtgtat cctcgatgag cctcaaaagt tctctcaccc cggataagaa 3300  
 acccttaagc aatgtgcaaa gtttgcatte tccactgaca taatgcaaaa taagatatca 3360  
 tcgatgacat agcaactcat gcatcatatc atgcctctct caacctatc attcctactc 3420  
 atctacataa gtatcttcag ctaaatgtta gaacataaac ccataagtca cgtttgatga 3480  
 gtattaggcg tgacacatga caaatcacag actcaagcaa gataaagcaa aatgatgtgt 3540  
 acataaaact ccagagctat atgtcatatt gcaaaaagag gagagcttat aagacaaggg 3600  
 atgactcaca aaaattcatt tgcctttcgt gtcaaaaaga ggagggtttt acattatcca 3660  
 tgtcatattg caaaagaaag agagaaagaa caacacaatg ctgcgtcaat tatacatatc 3720  
 tgtatgtcca tcattattca tccaccttct gtgtaccaca cttcatatat catgagtcac 3780  
 ttcatgtctg gacattaaca aactctatct taacatttag atgcaagagc ctttatctca 3840  
 ctataaatgc acgatgattt ctcatgtttt ctcaaaaaa gcattcagtt cattagtcct 3900  
 acaacaacga attcggcttc ccgggtacag ggtaaatttc tagtttttct ccttcatttt 3960  
 cttggtagg acccttttct ctttttattt ttttgagctt tgatctttct ttaaactgat 4020  
 ctatttttta attgattggg tatcgtgtaa atattacata gctttaactg ataactgat 4080  
 tactttattt cgtgtgtctt tgatcatctt gatagttaca gaaccgtcga ctctagagaa 4140  
 gccatttaaa tcgccgccac catgtctgtt gccttggtat gggttgtttc tccttggtgac 4200  
 gtctcaaacg ggacaggatt cttggtatcc gttcgtgagg gaaaccggat ttttgattcg 4260  
 tcggggcgta ggaatttggc gtgcaatgag agaatacaaga gaggaggtgg aaaacaaagg 4320  
 tggagttttg gttcttactt gggaggagca caaactggaa gtggacggaa attttctgta 4380  
 cgttctgcta tcgtggctac tccggctgga gaaatgacga tgatcatcaga acggatggta 4440  
 tatgatgtgg ttttgaggca ggcagccttg gtgaagagac agctgagatc gaccgatgag 4500  
 ttagatgtga agaaggatat acctattccg gggacttttg gcttggtgag tgaagcatat 4560  
 gataggtgta gtgaagtatg tgcagagtac gcaaagacgt ttacttagg aacgatgcta 4620  
 atgactccgg agagaagaaa ggctatctgg gcaatatacg tatggtgcag gagaacagac 4680  
 gaacttggtg atggtccgaa tgcacacac attactccgg cggccttaga taggtgggaa 4740  
 gacaggctag aagatgtttt cagtggacgg ccatttgaca tgctcgatgc tgctttgtcc 4800

-24-

gacacagttt ccaaatttcc agttgatatt cagccattca gagatatgat tgaaggaatg 4860  
 cgtatggact tgaggaagtc aagatacaga aactttgacg aactatacct atattgttat 4920  
 tacgttgctg gtacggtttg gttgatgagt gttccaatta tgggcatcgc acctgaatca 4980  
 aaggcaacaa cggagagcgt atataatgct gctttggcct tggggatcgc aaatcagctg 5040  
 accaacatac ttagagatgt tggagaagat gccagaagag gaagagtcta tttgcctcaa 5100  
 gatgaattag cacaggcagg tctatccgac gaagacatat ttgctggaag agtgaccgat 5160  
 aaatggagaa tcttcatgaa gaaacaaatt cagagggcaa gaaagttctt tgacgaggca 5220  
 gagaaaggag tgaccgaatt gagcgcagct agtagatggc ctgtgttggc atctctgctg 5280  
 ttgtaccgca ggatactgga cgagatcgaa gccaatgact acaacaactt cacaagaga 5340  
 gcttatgtga gcaaaccaaa gaagttgatt gcattaccta ttgcatatgc aaaatctctt 5400  
 gtgccttcta caagaacatg aaatcaggat tttatataaa tcaaggccaa tgaagccaat 5460  
 atacatttag aagaaaaaaaa acaagtgttt ataaagtaga attattgaag gggaggcttg 5520  
 gagtaactgg taaagttgtt gtcatgtgac tgggaagtca cgggttcaag ccttggaac 5580  
 agcctctggc agaaatgcaa ggtaagggtt cgtacaatat accgttaagg tggggtcctt 5640  
 cccagtacac cgcgcatagc gatagattta gtgcaccggg tcgccttttt tctaaagtag 5700  
 ggccatgcag gccgatcccc gatcgttcaa acatttggca ataaagtctt ttaagattga 5760  
 atcctgttgc cggctctgac atgattatca tataatttct gttgaattac gttaagcatg 5820  
 taataattaa catgtaatgc atgacgttat ttatgagatg ggtttttatg attagagtcc 5880  
 cgcaattata catttaatac gcgatagaaa acaaaatata gcgcgcaaac taggataaat 5940  
 tatcgcgcg cgtgtcatct atgttactag atcg 5974

&lt;210&gt; 8

&lt;211&gt; 5782

&lt;212&gt; DNA

&lt;213&gt; SYNTHETIC - 7650

&lt;400&gt; 8

gttaatcatg gtgtaggcaa ccaaataaaa acacaaaaat atgcacaagg cagtttgttg 60  
 tattctgtag tacagacaaa actaaaagta atgaaagaag atgtgggtgtt agaaaaggaa 120  
 acaatatcat gagtaatgtg tgagcattat gggaccacga aataaaaaga acattttgat 180  
 gagtctgtga tcctcgatga gcctcaaaag ttctctcacc ccggataaga aacccttaag 240

-25-

caatgtgcaa agtttgcatt ctccactgac ataatgcaaa ataagatata atcgatgaca	300
tagcaactca tgcacatata catgcctctc tcaacctatt cattcctact catctacata	360
agtatcttca gctaaatggt agaacataaa cccataagtc acgtttgatg agtattaggg	420
gtgacacatg acaaatcaca gactcaagca agataaagca aaatgatgtg tacataaaac	480
cccagagcta tatgtcatat tgcaaaaaga ggagagctta taagacaagg catgactcac	540
aaaaattcat ttgccttttcg tgtcaaaaag aggagggcctt tacattatcc atgtcatatt	600
gcaaaagaaa gagagaaaga acaacacaat gctgcgtcaa ttatacatat ctgtatgtcc	660
atcattatcc atccaccttt cgtgtaccac acttcatata tcatgagtca cttcatgtct	720
ggacattaac aaactctatc ttaacattta gatgcaagag cctttatctc actataaatg	780
cacgatgatt tctcattggt tctcacaaaa agcattcagt tcattagtcc tacaacaacg	840
aattcggcctt cccgggtaca gggtaaatTT ctagtttttc tcttcatttt tcttggttag	900
gacccttttc tctttttatt tttttgagct ttgatctttc tttaaactga tctatttttt	960
aattgattgg ttatcgtgta aatattacat agctttaact gataatctga ttactttatt	1020
tcgtgtgtct ttgatcatct tgatagttac agaaccgtcg actctagaga agccatttaa	1080
atcgccgccca ccatggcttc tatgatatac tcttcgcgtg tgacaacagt cagccgtgcc	1140
tctagggggc aatccgccgc agtggtcca ttcggcgggc tcaaatccat gactggatcc	1200
ccagtgaaga aggtcaacac tgacattact tccattacaa gcaatgggtg aagagtaaag	1260
tgcatggcgg ccgccaacc aactacggta attgggtgcag gcttcgggtg cctggcactg	1320
gcaattcgtc tacaagctgc ggggatcccc gtcttactgc ttgaacaacg tgataaaccc	1380
ggcggtcggg cttatgtcta cgaggatcag gggtttacct ttgatgcagg cccgacggtt	1440
atcacccgac ccagtgccat tgaagaactg tttgcaactg caggaaaaca gttaaaagag	1500
tatgtcgaac tgctgccggt tacgccgttt taccgcctgt gttgggagtc agggaaggtc	1560
tttaattacg ataacgatca aaccgggctc gaagcgcaga ttcagcagtt taatccccgc	1620
gatgtcgaag gttatcgtca gtttctggac tattcacgcg cgggtgttaa agaaggctat	1680
ctgaagctcg gtactgtccc ttttttatcg ttcagagaca tgcttcgcgc cgcacctcaa	1740
ctggcgaaac tgcaggcatg gagaagcgtt tacagtaagg ttgccagtta catcgaagat	1800
gaacatctgc gccaggcgtt ttctttccac tcgctgttggt tgggcggcaa tcccttcgcc	1860
acctcatcca tttatacggt gatacacgcg ctggagcgtg agtggggcgt ctggtttccg	1920
cgtggcgcca ccggcgcatt agttcagggg atgataaagc tgtttcagga tctgggtggc	1980
gaagtcgtgt taaacgccag agtcagccat atggaaacga caggaaacaa gattgaagcc	2040

-26-

gtgcatttag aggacggtcg caggttcctg acgcaagccg tcgcgtcaaa tgcagatgtg 2100  
 gttcatacct atcgcgacct gttaagccag caccctgccg cggttaagca gtccaacaaa 2160  
 ctgcagacta agcgcgatgag taactctctg tttgtgctct attttggttt gaatcaccat 2220  
 catgatcagc tcgcgcatca cacggtttgt ttcggccccg gttaccgcga gctgattgac 2280  
 gaaattttta atcatgatgg cctcgcagag gacttctcac tttatctgca cgcgccctgt 2340  
 gtcacggatt cgtcactggc gcctgaaggt tgcggcagtt actatgtgtt ggcgccggtg 2400  
 ccgcatttag gcaccgcgaa cctcgactgg acggttgagg ggccaaaact acgcgaccgt 2460  
 atttttgcgt accttgagca gcattacatg cctggcttac ggagtcagct ggtcacgcac 2520  
 cggatgttta cgcggttga ttttcgcgac cagcttaatg cctatcatgg ctcagccttt 2580  
 tctgtggagc ccgttcttac ccagagcgcc tggtttcggc cgcataaccg cgataaaacc 2640  
 attactaatc tctacctggt cggcgcaggc acgcatcccg gcgcaggcat tcttgccgtc 2700  
 atcggctcgg caaaagcgac agcaggtttg atgctggagg atctgatttg aggtacctcg 2760  
 acggccatgc aggccgatcc ccgatcgttc aaacatttgg caataaagtt tcttaagatt 2820  
 gaatcctgtt gccggtcttg cgatgattat catataatct ctgttgaatt acgttaagca 2880  
 tgtaataatt aacatgtaat gcatgacgtt atttatgaga tgggttttta tgattagagt 2940  
 cccgcaatta tacatttaat acgcgataga aaacaaaata tagcgcgcaa actaggataa 3000  
 attatcgcgc gcggtgtcat ctatgttact agatcgggcc ttaatcgcaa gcttgtaatt 3060  
 catggtgtag gcaaccctaa taaaacacca aaatatgcac aaggcagttt gttgtattct 3120  
 gtagtacaga caaaactaaa agtaatgaaa gaagatgtgg tgttagaaaa ggaaacaata 3180  
 tcatgagtaa tgtgtgagca ttatgggacc acgaaataaa aagaacattt tgatgagtcg 3240  
 tgtatcctcg atgagcctca aaagtctct caccocggat aagaaccct taagcaatgt 3300  
 gcaaagtttg cattctccac tgacataatg caaaataaga tatcatcgat gacatagcaa 3360  
 ctcatgcac atcatatgcc tctctcaacc tattcattcc tactcatcta cataagtac 3420  
 ttcagctaaa tgttagaaca taaaccata agtcacgttt gatgagtatt aggcgtgaca 3480  
 catgacaaat cacagactca agcaagataa agcaaatga tgtgtacata aaactccaga 3540  
 gctatatgtc atattgcaaa aagaggagag cttataagac aaggcatgac tcacaaaaat 3600  
 tcatttgcct ttcgtgtcaa aaagaggagg gctttacatt atccatgtca tattgcaaaa 3660  
 gaaagagaga aagaacaaca caatgctgcg tcaattatac atatctgtat gtccatcatt 3720  
 attcatccac ctttcgtgta ccacacttca tatatcatga gtcacttcat gtctggacat 3780  
 taacaaactc tatcttaaca tttagatgca agagccttta tctcactata aatgcacgat 3840

-27-

gattttctcat tgtttctcac aaaaagcatt cagttcatta gtcctacaac aacgaattcg 3900  
gcttccccggg tacagggtaa atttctagtt tttctccttc attttcttgg ttaggacct 3960  
tttctctttt tatttttttg agctttgatc tttctttaaa ctgatctatt ttttaattga 4020  
ttgggtatcg tgtaaatatt acatagcttt aactgataat ctgattactt tatttcgtgt 4080  
gtctttgatc atcttgatag ttacagaacc gtcgactcta gagaagccat ttaaatcgcc 4140  
gccaccatgt ctggtgcctt gttatgggtt gtttctcctt gtgacgtctc aaatgggaca 4200  
agtttcatgg aatcagtcg ggagggaaac cgttttttg attcatcgag gcataggaat 4260  
ttgggtgtcca atgagagaat caatagaggt ggtggaaagc aaactaataa tggacggaaa 4320  
ttttctgtac ggtctgctat tttggctact ccatctggag aacggacgat gacatcgga 4380  
cagatggtct atgatgtgg tttgaggcag gcagccttgg tgaagaggca actgagatct 4440  
accaatgagt tagaagtga gccggatata cctattccgg ggaatttggg cttggtgagt 4500  
gaagcatatg ataggtgtgg tgaagtatgt gcagagtatg caaagacgtt taacttagga 4560  
actatgctaa tgactcccga gagaagaagg gctatctggg caatatatgt atggtgcaga 4620  
agaacagatg aacttgttga tggcccaaac gcatcatata ttaccccggc agccttagat 4680  
agggtggaaa ataggctaga agatgttttc aatgggcggc catttgacat gctcgatggt 4740  
gctttgtccg atacagtttc taactttcca gttgatattc agccattcag agatatgatt 4800  
gaaggaatgc gtatggactt gagaaaatcg agatacaaaa acttcgacga actatacctt 4860  
tattgttatt atgttgctgg tacggttggg ttgatgagtgt ttccaattat gggatcgcc 4920  
cctgaatcaa aggcaacaac agagagcgta tataatgctg ctttggctct ggggatcgca 4980  
aatcaattaa ctaacatact cagagatggt ggagaagatg ccagaagagg aagagtctac 5040  
ttgcctcaag atgaattagc acaggcaggt ctatccgatg aagatatatt tgctggaagg 5100  
gtgaccgata aatggagaat ctttatgaag aaacaaatac atagggcaag aaagttcttt 5160  
gatgaggcag agaaaggcgt gacagaattg agctcagcta gtagattccc tgtatgggca 5220  
tctttggtct tgtaccgcaa aatactagat gagattgaag ccaatgacta caacaacttc 5280  
acaaagagag catatgtgag caaatcaaag aagttgattg cattacctat tgcatatgca 5340  
aaatctcttg tgccctctac aaaaactgcc tctcttcaaa gataaagcat gaaatgaaga 5400  
tatatatata tatatatata gcaatatata ttagaagaaa aaaaggaaga agaaatggtg 5460  
ttgtattgat ataaatgat atcataaata ttaggttgta gtaacattgg ccatgcaggc 5520  
cgatccccga tcgttcaaac atttggcaat aaagtttctt aagattgaat cctggtgccg 5580  
gtcttgcat gattatcata taatttctgt tgaattacgt taagcatgta ataattaaca 5640

-28-

tgtaatgcat gacgttattt atgagatggg tttttatgat tagagtcccg caattatata 5700  
 ttttaatacgc gatagaaaac aaaatatagc gcgcaaaacta ggataaatta tcgcgcgcgcg 5760  
 tgtcatctat gttactagat cg 5782

&lt;210&gt; 9

&lt;211&gt; 5551

&lt;212&gt; DNA

&lt;213&gt; SYNTHETIC

&lt;400&gt; .9

gttaatcatg gtgtaggcaa cccaaataaa acacaaaaat atgcacaagg cagtttggtg 60  
 tattctgtag tacagacaaa actaaaagta atgaaagaag atgtggtggt agaaaaggaa 120  
 acaatatcat gagtaatgtg tgagcattat gggaccacga aataaaaaaga acattttgat 180  
 gagtcgtgta tcttcgatga gcctcaaaag ttctctcacc ccggataaga aacccttaag 240  
 caatgtgcaa agtttgcatt ctccactgac ataatgcaaa ataagatatc atcgatgaca 300  
 tagcaactca tgcacatcat catgcctctc tcaacctatt cattcctact catctacata 360  
 agtatcttca gctaaatggt agaacataaa cccataagtc acgtttgatg agtattaggg 420  
 gtgacacatg acaaatcaca gactcaagca agataaagca aaatgatgtg tacataaaac 480  
 tccagagcta tatgtcatat tgcaaaaaga ggagagctta taagacaagg catgactcac 540  
 aaaaattcat ttgcctttcg tgtcaaaaag aggagggctt tacattatcc atgtcatatt 600  
 gcaaaagaaa gagagaaaga acaacacaat gctgcgtcaa ttatacatat ctgtatgtcc 660  
 atcattatcc atccaccttt cgtgtaccac acttcatata tcatgagtca cttcatgtct 720  
 ggacattaac aaactctatc ttaacattta gatgcaagag cctttatctc actataaatg 780  
 cacgatgatt tctcattggt tctcacaaaa agcattcagt tcattagtcc tacaacaacg 840  
 aattcggctt cccgggtaca gggtaaattt ctagttttct tccttcattt tcttggttag 900  
 gacccttttc tctttttatt tttttgagct ttgatctttc tttaaactga tctatttttt 960  
 aattgattgg ttatcgtgta aatattacat agctttaact gataatctga ttactttatt 1020  
 tcgtgtgtct ttgatcatct tgatagttac agaaccgtcg actctagaga agccatttaa 1080  
 atcgccgcca ccattggttc tatgatatcc tcttcgcgtg tgacaacagt cagccgtgcc 1140  
 tctagggggc aatccgcccgc agtgggtcca ttccgcccgc tcaaatccat gactggattc 1200  
 ccagtgaaga aggtcaaacac tgacattact tccattacaa gcaatggtgg aagagtaaag 1260



-29-

tgcattggcgg cgcgcaaac aactacggta attgggtgcag gcttcgggtg cctggcactg	1320
gcaattcgtc tacaagctgc ggggatcccc gtcttactgc ttgaacaacg tgataaaccc	1380
ggcggtcggg cttatgtcta cgaggatcag ggggtttacct ttgatgcagg cccgacgggt	1440
atcaccgatc ccagtgccat tgaagaactg ttgactgagg caggaaaaca gttaaaagag	1500
tatgtcgaac tgctgccggg tacgccgttt taccgcctgt gttgggagtc agggaggtc	1560
tttaattacg ataacgatca aaccgggtc gaagcgcaga ttcagcagtt taatccccgc	1620
gatgtcgaag gttatcgtca gtttctggac tattcacgcg cgggtgtttaa agaaggctat	1680
ctgaagctcg gtactgtccc ttttttatcg ttcagagaca tgcttcgcgc cgcacctcaa	1740
ctggcgaaac tgcaggcatg gagaagcgtt tacagtaagg ttgccagtta catcgaagat	1800
gaacatctgc gccaggcgtt ttctttccac tcgctgttgg tgggcggcaa tcccttcgcc	1860
acctcatcca tttatacgtt gatacacgcg ctggagcgtg agtggggcgt ctgggttccg	1920
cgtggcggca ccggcgcat agttcagggg atgataaagc tgtttcagga tctgggtggc	1980
gaagtcgtgt taaacgccag agtcagccat atggaaacga caggaaacaa gattgaagcc	2040
gtgcatttag aggacggtcg caggttcctg acgcaagccg tcgcgtcaaa tgcagatgtg	2100
gttcataacct atcgcgacct gttaagccag caccctgccg cggttaagca gtccaacaaa	2160
ctgcagacta agcgcagtag taactctctg tttgtgctct attttggttt gaatcaccat	2220
catgatcagc tcgcgcacat cacggtttgt ttccggccgc gttaccgcga gctgattgac	2280
gaaattttta atcatgatgg cctcgcagag gacttctcac tttatctgca cgcgccctgt	2340
gtcacggatt cgtcactggc gcctgaaggc tgcggcagtt actatgtgtt ggccgggtg	2400
ccgcatttag gcacgcgaa cctcgcactg acgggttgagg ggccaaaact acgcgaccgt	2460
atTTTTgcgt accttgagca gcattacatg cctggcttac ggagtcagct ggtcacgcac	2520
cggatgttta cgcggtttga ttttcgcgac cagcttaatg cctatcatgg ctacgccttt	2580
tctgtggagc ccgttcttac ccagagcgcc tggtttcggc cgcataaccg cgataaaacc	2640
attactaatc tctacctggc cggcgcaggc acgcaccccg gcgcaggcat tcttggcgtc	2700
atcggctcgg caaaagcgac agcaggtttg atgctggagg atctgatttg aggtacctcg	2760
acggccatgc aggcgatcc ccgatcgtt aaacatttgg caataaagtt tcttaagatt	2820
gaatcctgtt gccggtcttg cgatgattat catataattt ctgttgaatt acgttaagca	2880
tgtaataatt aacatgtaat gcatgacgtt atttatgaga tgggttttta tgattagagt	2940
cccgcaatta tacatttaac acgcgataga aaacaaaata tagcgcgcaa actaggataa	3000
attatcgcgc gcggtgtcat ctatgttact agatcgggccc ttaatcgcaa gcttggtta	3060

-30-

catggtgtag gcaacccaaa taaaacacca aaatatgcac aaggcagttt gttgtattct 3120  
gtagtacaga caaaactaaa agtaatgaaa gaagatgtgg tgtagaaaaa ggaaacaata 3180  
tcatgagtaa tgtgtgagca ttatgggacc acgaaataaa aagaacattt tgatgagtcg 3240  
tgtatcctcg atgagcctca aaagtctctc caccgccgat aagaaaccct taagcaatgt 3300  
gcaaagtttg cattctccac tgacataatg caaaataaga tatcatcgat gacatagcaa 3360  
ctcatgcac atcatcgcc tctctcaacc tattcattcc tactcatcta cataagtatc 3420  
ttcagctaaa tgtagaaca taaaccata agtcacgttt gatgagtatt aggcgtgaca 3480  
catgacaaat cacagactca agcaagataa agcaaatga tgtgtacata aaactccaga 3540  
gctatatgtc atattgcaaa aagaggagag cttataagac aaggcatgac tcacaaaaat 3600  
tcatttgctt ttcgtgtcaa aaagaggagg gctttacatt atccatgtca tattgcaaaa 3660  
gaaagagaga aagaacaaca caatgctgcg tcaattatac atatctgtat gtccatcatt 3720  
attcatccac ctttcgtgta ccacacttca tatatcatga gtcacttcat gtctggacat 3780  
taacaaactc tatcttaaca tttagatgca agagccttta tctcactata aatgcacgat 3840  
gatttctcat tgtttctcac aaaaagcatt cagttcatta gtcctacaac aacgaattcg 3900  
gcttcccggt tacagggtaa atttctagtt tttctccttc attttcttgg ttaggacctt 3960  
tttctctttt tatttttttg agctttgatc tttctttaa ctgatctatt ttttaattga 4020  
ttggttatcg tgtaaatatt acatagcttt aactgataat ctgattactt tatttcgtgt 4080  
gtctttgatc atcttgatag ttacagaacc gtcgactcta gagaagccat ttaaategcc 4140  
gccaccatgg cttctatgat atcctcttcc gctgtgacaa cagtcagccg tgcctctagg 4200  
gggcaatccg ccgcagtggc tccattcggc ggctcaaat ccatgactgg attcccagtg 4260  
aagaaggtca aactgacat tacttcatt acaagcaatg gtggaagagt aaagtgcag 4320  
gcagttggct cgaaaagttt tgcgacagcc tcaaagttat ttgatgcaa aaccggcg 4380  
agcgtactga tgctctacgc ctggtgccgc cattgtgacg atgttattga cgatcagacg 4440  
ctgggcttct aggccgggca gcctgcctta caaacgccc aacaacgtct gatgcaactt 4500  
gagatgaaaa cgcgccaggc ctatgcagga tcgcagatgc acgaaccggc gtttgccgct 4560  
tttcaggaag tggctatggc tcatgatatc gccccggctt acgcgtttga tcatctggaa 4620  
ggcttcgcga tggatgtacg cgaagcgcaa tacagccaac tggatgatac gctgcgctat 4680  
tgctatcacg ttgcaggcgt tgcggcttg atgatggcg aaatcatggg cgtgcgggat 4740  
aacgccacgc tggaccggc ctgtgacctt gggctggcat ttcagttgac caatattgct 4800  
cgcatattg tggacgatgc gcatgcgggc cgctgttatc tgccggcaag ctggctggag 4860

-31-

catgaagggtc tgaacaaaga gaattatgcg gcacctgaaa accgtcagggc gctgagccgt 4920  
 atcgcccgac gtttggtgca ggaagcagaa cttactatt tgtctgccac agccggcctg 4980  
 gcagggttgc ccctgcgttc cgcctgggca atcgctacgg cgaagcaggt ttaccggaaa 5040  
 atagggtgtca aagttgaaca ggccgggtcag caagcctggg atcagcggca gtcaacgacc 5100  
 acgcccga aaattaacgct gctgctggcc gcctctgggtc aggcccttac ttcccggatg 5160  
 cgggctcatc ctccccgccc tgcgcatctc tggcagcgcc cgctctaggg atccgttaag 5220  
 ggcgaattcc agcacttg cgcccggttac tagtggtacc gagctcggtc cctcgacggc 5280  
 catgcaggcc gatccccgat cgttcaaaca ttggcaata aagtttctta agattgaatc 5340  
 ctgttgccgg tcttgcatg attatcatat aatttctgtt gaattacgtt aagcatgtaa 5400  
 taattaacat gtaatgcatg acgttattta tgagatgggt ttttatgatt agagtccgc 5460  
 aattatacat ttaatacgcg atagaaaaca aaatatagcg cgcaaactag gataaattat 5520  
 cgcgcgcggt gtcactatg ttactagatc g 5551

&lt;210&gt; 10

&lt;211&gt; 1233

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;400&gt; 10

atggccatca tactcgtacg agcagcgctc cgggggtctt ccgccgccga cagcatcagc 60  
 caccagggga ctctccagtg ctccaccctg ctcaagacga agaggccggc ggcgcgggcg 120  
 tggatgccct gctcgtctct tggcctccac ccgtgggagg ctggccgtcc ctccccgc 180  
 gtctactcca gcctgccgt caaccggcg ggagaggccg tcgtctcgtc cgagcagaag 240  
 gtctacgacg tcgtgctcaa gcaggccgca ttgctcaaac gccagctgcg cagccgggtc 300  
 ctcgacgcca ggccccagga catggacatg ccacgcaacg ggctcaagga agcctacgac 360  
 cgctgcggcg agatctgtga ggagtatgcc aagacgtttt acctcggaac tatgttgatg 420  
 acagaggagc ggcgccgccc catatgggcc atctatgtgt ggtgtaggag gacagatgag 480  
 cttgtagatg ggccaaacgc caactacatt acaccaacag ctttgaccg gtgggagaag 540  
 agacttgagg atctgttcac gggacgtctt tacgacatgc ttgatgccgc tctctctgat 600  
 accatctcaa ggttccccat agacattcag ccattcaggg acatgattga agggatgagg 660  
 agtgatctta ggaagacaag gtataacaac ttcgacgagc tctacatgta ctgctactat 720

-32-

gttgctggaa ctgtcgggtt aatgagcgta cctgtgatgg gcatcgcaac cgagtctaaa 780  
gcaacaactg aaagcgata cagtgtgcc ttggctctgg gaattgcgaa ccaactcacg 840  
aacatactcc gggatgttgg agaggatgct agaagaggaa ggatatattt accacaagat 900  
gagcttgcac aggcagggtt ctctgatgag gacatcttca aaggggtcgt cacgaaccgg 960  
tggagaaact tcatgaagag gcagatcaag agggccagga tgttttttga ggaggcagag 1020  
agaggggtaa ctgagctctc acaggctagc agatggccag tatgggcttc cctgttgttg 1080  
tacaggcaga tcctggatga gatcgaagcc aacgactaca acaacttcac gaagagggcg 1140  
tatgttggtg aaggaagaa gttgctagca ctctctgtgg catatggaaa atcgctactg 1200  
ctcccatgtt cattgagaaa tggccagacc tag 1233

&lt;210&gt; 11

&lt;211&gt; 1233

&lt;212&gt; DNA

&lt;213&gt; Zea mays

<400> 11  
atggccatca tactcgtacg agcagcgctg cgggggtctt cgcgcgcga cagcatcagc 60  
caccagggga ctctccagtg ctccaccctg ctcaagacga agaggccggc ggcgcggcgg 120  
tggatgccct gctcgtctct tggcctccac ccgtgggagg ctggcgcgtc ctcccccgcc 180  
gtctactcca gcctgccctg caaccggcg ggagaggccg tcgtctcgtc cgagcagaag 240  
gtctacgacg tcgtgtctaa gcaggccgca ttgtctaaac gccagctgcg cagcgggtc 300  
ctcgacgcca ggcccagga catggacatg ccacgcaacg ggctcaagga agcctacgac 360  
cgctgcggcg agatctgtga ggagtatgcc aagacgtttt acctcggaac tatgttgatg 420  
acagaggagc ggcgcgcgc catatgggcc atctatgtgt ggtgtaggag gacagatgag 480  
cttgtagatg ggccaaacgc caactacatt acaccaacag ctttgaccg gtgggagaag 540  
agacttgagg atctgttcac gggacgtctt tacgacatgc ttgatgccgc tctctctgat 600  
accatctcaa ggttccccat agacattcag ccattcaggg acatgattga agggatgagg 660  
agtgatctta ggaagacaag gtataacaac ttgcacgagc tctacatgta ctgtactat 720  
gttgctggaa ctgtcgggtt aatgagcgta cctgtgatgg gcatcgcaac cgagtctaaa 780  
gcaacaactg aaagcgata cagtgtgcc ttggctctgg gaattgcgaa ccaactcacg 840  
aacatactcc gggatgttgg agaggatgct agaagaggaa ggatatattt accacaagat 900

-33-

gagcttgac aggcagggct ctctgatgag gacatcttca aaggggtcgt cacgaaccgg 960  
 tggagaaact tcatgaagag gcagatcaag agggccagga tgttttttga ggaggcagag 1020  
 agaggggtaa atgagctctc acaggctagc agatggccag tatgggcttc cctgttggtg 1080  
 tacaggcaga tcctggatga gatcgaagcc aacgactaca acaacttcac gaagagggcg 1140  
 tatgttggtg aaggggaagaa gttgctagca cttcctgtgg catatggaaa atcgctactg 1200  
 ctcccatggt cattgagaaa tggccagacc tag 1233

&lt;210&gt; 12

&lt;211&gt; 1233

&lt;212&gt; DNA

&lt;213&gt; Zea mays

<400> 12  
 atggccatca tactcgtagc agcagcgtag ccggggctct ccgccgccga cagcatcagc 60  
 caccagggga ctctccagtg ctccaccctg ctcaagacga agaggccggc ggcgcgccgg 120  
 tggatgccct gctcgctcct tggcctccac ccgtgggagg ctggccgtcc ctccccgcc 180  
 gtctactcca gcctcgccgt caaccggcg ggagaggccg tcgtctcgtc cgagcagaag 240  
 gtctacgacg tcgtgctcaa gcaggccgca ttgctcaaac gccagctgag cacgccggtc 300  
 ctcgacgcca ggccccagga catggacatg ccacgcaacg ggctcaagga agcctacgac 360  
 cgctgcggcg agatctgtga ggagtatgcc aagacgtttt acctcggaac tatgttgatg 420  
 acagaggagc ggcgccgcgc catatgggcc atctatgtgt ggtgtaggag gacagatgag 480  
 cttgtagatg ggccaaacgc caactacatt acaccaacag ctttggaacc gtgggagaag 540  
 agacttgagg atctgttcac gggacgtcct tacgacatgc ttgatgccgc tctctctgat 600  
 accatctcaa ggttcccat agacattcag ccattcaggg acatgattga agggatgagg 660  
 agtgatctta ggaagacaag gtataacaac ttcgacgagc tctacatgta ctgctactat 720  
 gttgctggaa ctgtcgggtt aatgagcgta ccagtgatgg gcatcgcatc cgagtctaaa 780  
 gcaacaactg aaagcgtgta cagtgtgccc ttggctctcg gaattgcgaa ccaactcacg 840  
 aacatactcc gggatgttg agaggatgct agacgaggaa ggatatattt accacaagat 900  
 gagcttgac aggcagggct ctctgatgag gacatcttca aaggggtcgt cacgaaccgg 960  
 tggagaaact tcatgaagag gcagatcaag agggccagga tgttttttga ggaggcagag 1020  
 agaggggtaa ctgagctctc acaggctagc agatggccag tatgggcttc cctgttggtg 1080

-34-

tacaggcaga tcctggatga gatcgaagcc aacgactaca acaacttcac gaagagggcg 1140  
 tatgttggtg aaggggaagaa gttgctagca cttcctgtgg catatggaaa atcgctactg 1200  
 ctcccatggt cattgagaaa tggccagacc tag 1233

&lt;210&gt; 13

&lt;211&gt; 1263

&lt;212&gt; DNA

&lt;213&gt; Oryza sp.

&lt;400&gt; 13

atggcgccca tcacgtcct acgttcagcg tctcttcgg gcctctccga cgcctcgcc 60  
 cgggacgctg ctgccgtcca acatgtctgc tcctcctacc tgcccaacaa caaggagaag 120  
 aagaggaggt ggatcctctg ctgctcaag tacgcctgcc ttggcgctga ccctgccccg 180  
 ggcgagattg cccggacctc gccggtgtac tccagcctca cgtcaccctc tgctggagag 240  
 gccgtcatct cctcggagca gaaggtgtac gacgtcgtcc tcaagcaggc agcattgctc 300  
 aaacgccacc tgcgccaca accacacacc attcccatcg ttcccaagga cctggacctg 360  
 ccaagaaacg gcctcaagca ggcctatcat cgctgcggag agatctgcga ggagtatgcc 420  
 aagacctttt accttgaac tatgctcatg acggaggacc gacggcgcg catatgggcc 480  
 atctatgtgt ggtgtaggag gacagatgag cttgtagatg gaccaaagtc ctgcacatc 540  
 acaccgtcag ccctggaccg gtgggagaag aggcttgatg atctcttcac cggacgcccc 600  
 tacgacatgc ttgatgctgc actttctgat accatctcca agtttcttat agatattcag 660  
 cctttcaggg acatgataga agggatgcgg tcagacctca gaaagactag atacaagaac 720  
 ttgacgagc tctacatgta ctgctactat gttgctggaa ctgtggggct aatgagtgtt 780  
 cctgtgatgg gtattgcacc cgagtcgaag gcaacaactg aaagtgtgta cagtgtgtct 840  
 ttggctctcg gcattgcaaa ccagctcaca aatatactcc gtgacgttgg agaggacgcg 900  
 agaagagggg ggatatatctt accacaagat gaacttcag aggcagggt ctctgatgag 960  
 gacatcttca atggcgttgt gactaacaac tggagaagct tcatgaagag acagatcaag 1020  
 agagctagga tgttttttga ggaggcagag agaggggtga ccgagctcag ccaggcaagc 1080  
 cgggtggcgg tctgggcgtc tctgttgta taccggcaaa tccttgacga gatagaagca 1140  
 aacgattaca acaacttcac aaagagggcg tacgttggga aggcgaagaa attgctagcg 1200  
 cttccagttg catatggtag atcattgctg atgccctact cactgagaaa tagccagaag 1260

- 26 -

thus modified sequence to hybridise to a sequence in the Sequence Listing under the conditions described above.

The present invention further provides a plant which comprises a polynucleotide or polynucleotide sequence as described above. In a particular embodiment said plant is  
5 a rice or a maize plant.

In a further aspect of the present invention there is provided a polynucleotide or a polynucleotide sequence as described above wherein the promoter(s) are tissue preferred and/or organ preferred. In a particular embodiment the promoter(s) provide for preferential expression in fruit. In a further embodiment said promoter(s) provide for  
10 high expression in fruit. In a still further embodiment said fruit is a banana fruit.

In a further aspect of the present invention there is provided a polynucleotide comprising:  
(a) a region which comprises as operably linked components (i) a promoter which provides for fruit preferred expression; and (ii) a nucleotide sequence derived from a bacterium which sequence encodes a carotene desaturase; and (iii) a transcription  
15 termination region; and (b) a further region which comprises as operably linked components (i) a promoter which provides for fruit preferred expression; and (ii) a nucleotide sequence encoding a phytoene synthase which sequence is derived from maize (*Zea sp.*) or rice (*Orzya sp.*); and (iii) a transcription termination region.

In a still further embodiment there is provided a method for increasing the carotenoid  
20 content of fruit comprising inserting into plant material a polynucleotide comprising:  
(a) a region which comprises as operably linked components (i) a promoter which provides for fruit preferred expression; and (ii) a nucleotide sequence derived from a bacterium which sequence encodes a carotene desaturase; and (iii) a transcription termination region; and (b) a further region which comprises as operably linked  
25 components (i) a promoter which provides for fruit preferred expression; and (ii) a nucleotide sequence encoding a phytoene synthase which sequence is derived from maize (*Zea sp.*) or rice (*Orzya sp.*); and (iii) a transcription termination region and regenerating a fruit-containing plant from said material and identifying the fruit which contain carotenoids at levels greater than those of control like-fruit. The present invention still  
30 further provides polynucleotides which comprise the phytoene synthase encoding sequences and carotene desaturase encoding sequences mentioned above which sequences are operably linked to promoters which provide for fruit preferred or tissue or

organ preferred expression. Such suitable promoters may be identified by the person skilled in the art.

In a further aspect of the present invention there is provided the use of a polynucleotide or a polynucleotide sequence as described above in a method for the production of plants which are resistant and/or tolerant to a herbicide.

In a still further aspect of the present invention there is provided a method for the production of a plant that is resistant and/or tolerant to a herbicide comprising inserting into plant material a polynucleotide or a polynucleotide sequence as described above and regenerating a morphologically normal plant from said material. The herbicide resistance and/or tolerance of the plant containing the polynucleotide or polynucleotide sequence of the invention can be compared to a control like-plant. The term control like-plant relates to plants which are substantially similar to those according to the invention but which control like-plant does not contain the polynucleotides or polynucleotide sequences according to the invention. Typically, a control like-plant will comprise a plant of the same or similar plant species which control like-plant is a native plant or which has not been transformed.

In a further aspect of the present invention there is provided a polynucleotide comprising the sequence depicted as SEQ ID NO: 13. In a particular embodiment there is provided a polynucleotide which consists of the sequence depicted as SEQ ID NO: 13. The present invention still further provides a polynucleotide which encodes the protein depicted as SEQ ID NO: 14. The present invention still further provides a polynucleotide sequence which has at least 87% identity to the sequence depicted as SEQ ID NO: 13 wherein said sequence still encodes a phytoene synthase. The present invention still further provides a polynucleotide sequence which has at least 90% identity to the sequence depicted as SEQ ID NO: 13 wherein said sequence still encodes a phytoene synthase. The present invention still further provides a polynucleotide sequence which has at least 91% identity to the sequence depicted as SEQ ID NO: 13 wherein said sequence still encodes a phytoene synthase. The present invention still further provides a polynucleotide sequence which has at least 92% identity to the sequence depicted as SEQ ID NO: 13 wherein said sequence still encodes a phytoene synthase. The present invention still further provides a polynucleotide sequence which has at least 93% identity to the sequence depicted as SEQ ID NO: 13 wherein said sequence still encodes a phytoene synthase. The present invention still further provides a polynucleotide



sequence which has at least 94% identity to the sequence depicted as SEQ ID NO: 13 wherein said sequence still encodes a phytoene synthase. The present invention still further provides a polynucleotide sequence which has at least 95% identity to the sequence depicted as SEQ ID NO: 13 wherein said sequence still encodes a phytoene  
5 synthase. The present invention still further provides a polynucleotide sequence which has at least 96% identity to the sequence depicted as SEQ ID NO: 13 wherein said sequence still encodes a phytoene synthase. The present invention still further provides a polynucleotide sequence which has at least 97% identity to the sequence depicted as SEQ ID NO: 13 wherein said sequence still encodes a phytoene synthase. The present  
10 invention still further provides a polynucleotide sequence which has at least 98% identity to the sequence depicted as SEQ ID NO: 13 wherein said sequence still encodes a phytoene synthase. The present invention still further provides a polynucleotide sequence which has at least 99% identity to the sequence depicted as SEQ ID NO: 13 wherein said sequence still encodes a phytoene synthase.

15 The present invention still further provides a protein having the sequence depicted as SEQ ID NO: 14 or a variant which has at least 82% identity to SEQ ID NO: 14 wherein said variant still provides for phytoene synthase activity. In a further embodiment said variant has at least 85% identity to SEQ ID NO: 14 wherein said variant still provides for phytoene synthase activity. In a still further embodiment said  
20 variant has at least 90% identity to SEQ ID NO: 14 wherein said variant still provides for phytoene synthase activity. In a still further embodiment said variant has at least 91% identity to SEQ ID NO: 14 wherein said variant still provides for phytoene synthase activity. In a still further embodiment said variant has at least 92% identity to SEQ ID NO: 14 wherein said variant still provides for phytoene synthase activity. In a still further  
25 embodiment said variant has at least 93% identity to SEQ ID NO: 14 wherein said variant still provides for phytoene synthase activity. In a still further embodiment said variant has at least 94% identity to SEQ ID NO: 14 wherein said variant still provides for phytoene synthase activity. In a still further embodiment said variant has at least 95% identity to SEQ ID NO: 14 wherein said variant still provides for phytoene synthase  
30 activity. In a still further embodiment said variant has at least 96% identity to SEQ ID NO: 14 wherein said variant still provides for phytoene synthase activity. In a still further embodiment said variant has at least 97% identity to SEQ ID NO: 14 wherein said variant still provides for phytoene synthase activity. In a still further embodiment said

variant has at least 98% identity to SEQ ID NO: 14 wherein said variant still provides for phytoene synthase activity. In a still further embodiment said variant has at least 99% identity to SEQ ID NO: 14 wherein said variant still provides for phytoene synthase activity. In a still further embodiment there is provided a polynucleotide which encodes said variant. By phytoene synthase activity it is meant that the variant protein has the same or a similar function to the protein depicted as SEQ ID NO: 14. The percentage of sequence identity for proteins is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the amino acid sequence in the comparison window may comprise additions or deletions (i.e. gaps) as compared to the initial reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical amino acid residue occurs in both sequences to yield the number of match positions, dividing the number of match positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. Optimal alignment of sequences for comparison may also be conducted by computerised implementations of known algorithms such as Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402. There are also algorithms available to the person skilled in the art that enable a calculation of the percentage sequence identity between polynucleotide sequences. The variant may differ from the protein depicted as SEQ ID NO: 14 in particular by conservative substitutions. Such conservative substitutions are described above.

The present invention will now be described by way of the following non-limiting examples with reference to the following Figures and Sequence Listing of which:

- SEQ ID NO: 1 = 12423 = Glu-Cat-SSU-crtI-Nos-Glu-Cat-Psy (Maize-gb)-nos
- 30 SEQ ID NO: 2 = 12421 = Glu-Cat-SSU-crtI-Nos-Glu-Cat-Psy (Maize-E1B)-nos
- SEQ ID NO: 3 = 12422 = Glu-SSU-crtI-Nos-Glu-Psy (Maize-E1B)-nos
- SEQ ID NO: 4 = 12424 = Glu-SSU-crtI-Nos-Glu-Psy (Maize-gb)-nos
- SEQ ID NO: 5 = Glu-Cat-SSU-crtI-Nos-Glu-Cat-Psy (Maize)-nos

- SEQ ID NO: 6 = 11586 = Glu-Cat-SSU-crtI-Nos-Glu-Cat-Psy (Rice)-nos  
 SEQ ID NO: 7 = 7651 = Glu-Cat-SSU-crtI-Nos-Glu-Cat-Psy (Pepper)-nos  
 SEQ ID NO: 8 = 7650 = Glu-Cat-SSU-crtI-Nos-Glu-Cat-Psy (Tomato)-nos  
 SEQ ID NO: 9 = Glu-Cat-SSU-crtI-Nos-Glu-Cat-SSU-Psy (crtB)-nos  
 5 SEQ ID NO: 10 = Phytoene synthase gb (Maize)  
 SEQ ID NO: 11 = Phytoene synthase (Maize) from SEQ ID NO 5 above  
 SEQ ID NO: 12 = Phytoene synthase E1B (Maize)  
 SEQ ID NO: 13 = Phytoene synthase (Rice)  
 SEQ ID NO: 14 = Phytoene synthase (Rice) PROTEIN  
 10 SEQ ID NO: 15 = Phytoene synthase (Pepper) .  
 SEQ ID NO: 16 = Phytoene synthase (Tomato)  
 SEQ ID NO: 17 = Phytoene synthase (*Erwinia* crtB)  
 SEQ ID NO: 18 = Carotene desaturase (*Erwinia* crtI) used in SEQ ID NOS: 1-4  
 SEQ ID NO: 19 = Carotene desaturase (*Erwinia* crtI)  
 15 SEQ ID NO: 20 = Glutelin seed preferred promoter  
 SEQ ID NO: 21 = Prolamin seed preferred promoter  
 SEQ ID NO: 22 = Intron from catalase gene  
 SEQ ID NO: 23 = Transit peptide (Small sub-unit Rubisco)  
 SEQ ID NO: 24 = Transcription termination region from nopaline synthase gene  
 20 SEQ ID NO: 25 = Transcription termination region from 35S CaMV  
 SEQ ID NO: 26 = Transcription termination region from proteinase inhibitor from potato  
 SEQ ID NO: 27 = Carotene desaturase (Tomato)  
 SEQ ID NO: 28 = Carotene desaturase (Pepper)  
 SEQ ID NO: 29 = Carotene desaturase (Maize)  
 25 SEQ ID NO: 30 = Carotene desaturase (Rice)  
 SEQ ID NO: 31 = Zeta-carotene desaturase (Tomato)  
 SEQ ID NO: 32 = Zeta-carotene desaturase (Pepper)  
 SEQ ID NO: 33 = Zeta-carotene desaturase (Maize)  
 SEQ ID NO: 34 = Zeta-carotene desaturase (Rice)  
 30 SEQ ID NOS: 35 to 38 = Primers

Glu = Glutelin seed preferred promoter

Pro = Prolamin seed preferred promoter

Cat = Intron from catalase gene

SSU = Transit peptide (Small sub-unit Rubisco)

CrtI = Carotene desaturase from *Erwinia*

Pds = Carotene desaturase (source indicated in parenthesis)

5 Psy = Phytoene synthase (source indicated in parenthesis)

Zds = Zeta-carotene desaturase

Pds = Phytoene desaturase

Nos = Transcription termination region from nopaline synthase gene

35S term = Transcription termination region from 35S CaMV

10 PotP1-II term = Transcription termination region from proteinase inhibitor from potato

FIGURE 1 = Part of carotenoid biosynthesis pathway starting from GGPP  
(Geranylgeranyl diphosphate).

FIGURE 2 = Construct pPRP0117.

15

## EXAMPLES

General molecular biology methods are carried out according to Sambrook et al  
(1989) 'Molecular cloning: A laboratory Manual, 2nd Edition. Cold Spring Harbour Lab.  
20 Press.

### 1.0 Construction of plant binary vectors

References for the gene sequences below are given as they are listed on the EMBL  
database. This database is maintained and distributed by the European Bioinformatics  
25 Institute (Patricia Rodriguez-Tomé, Peter J. Stoeckl, Graham N. Cameron and Tomas P.  
Flores, "The European Bioinformatics Institute (EBI) databases", Nucleic Acids Res.  
24:(6-13), 1996, www.ebi.ac.uk.)

A pUC based vector pPRP0117 (Figure 2) was used for cloning of all the plant  
30 transformation vectors. This contains the nucleotides -806 to +33 of the rice glutelin  
gene as a promoter (Y00687), the first intron of the catalase-1 gene from castor bean  
altered to remove the ATG sequences, a *gus* coding region and a *nos* terminator. The *gus*  
coding sequence was removed by digestion of pPRP0117 with NcoI and SfiI and

replaced by the coding regions of the carotenoid phytoene synthase genes or phytoene desaturase genes, or removed as a *gus::nos* cassette by digestion of pPRP0117 with NcoI and PacI and replaced by a carotenoid coding region/ *nos* terminator fusion as described below.

#### 1.1 Construction of the pPRP0117 +*crtI* vector

A cassette of the signal peptide of the small subunit of pea ribulose biphosphate carboxylase (*SSU*) (X00806) fused to the bacterial phytoene desaturase *CrtI* (D90087) and a *nos* terminator was cloned into the NcoI and PacI sites of pPRP0117. The *crtI* sequence in constructs 7651, 7650 and 11586 had 9 nucleotides extra (3 alanines) inserted after the first ATG in order to incorporate a NotI restriction site for cloning purposes.

#### 1.2 Construction of the pJH0104HygCrtI binary vector

The SgfI Gt::*intron*::*SSUcrtI*::*nos* cassette was cloned into the PacI site of the binary vector pJH0104+Hyg (which contains a hygromycin resistance gene for antibiotic selection) to give the construct pJH0104HygSSUCrtI.

#### 1.3 Pepper *psy* + *crtI* construct 7651

The catalase intron and pepper *psy* (X68017) were separately amplified by PCR using primers that overlap both sequences, and then fused by recombinant PCR, and cloned into the EcoRI and SfiI sites of pPRP0117. The Gt::*intron*::pepper *psy*::*nos* cassette was recovered with SgfI digestion and cloned into the PacI site of pJH0104HygCrtI to give the construct 7651.

#### 1.4 Tomato *psy* + *crtI* construct 7650

The catalase intron and tomato *psy* (Y00521) were separately amplified by PCR using primers that overlap both sequences, and then fused by recombinant PCR and cloned into the EcoRI and SfiI sites of pPRP0117. The Gt::*intron*::tomato *psy*::*nos* cassette was recovered by SgfI and cloned into the PacI site of pJH0104HygCrtI to give the construct 7650.

### 1.5 Rice psy + crtI construct 11586

PolyA mRNA was extracted from rice leaves (Asanohikari). First strand synthesis of Rice *psy* cDNA was synthesized using the antisense primer 5' cgtcggcctgcatggccctacttctggctatttctcagtg 3' (SEQ ID NO: 35) and cDNA was then obtained by PCR amplification with this antisense primer and the sense primer 5'ctgtccatggcggccatcacgctcct 3' (SEQ ID NO: 36). This was digested with NcoI and SfiI and cloned into pPRP0117. The Gt::intron::rice psy::nos fragment was transferred to the binary vector pJH0104Hyg. A HindIII/PacI Gt::intron::SSUcrtI::nos cassette was blunt ended and cloned into the PmeI site of the pJH0104HygRicepsy vector to give the construct 11586.

### 1.6 Maize psy (E1B) + crtI construct 12421

PolyA mRNA was extracted from maize leaves. First strand synthesis of maize *psy* (sequence designated "E1B") cDNA was synthesized using the antisense primer 5' cgatggcctgcatggccctaggtctggccatttctcaatg 3' (SEQ ID NO: 37) and cDNA was then obtained by PCR amplification with this antisense primer and the sense primer 5' taggataagatagcaaatccatggccatcata 3' (SEQ ID NO: 38). This was digested with NcoI and SfiI and cloned into the NcoI and SfiI sites of a pPRP0117 based vector. The Gt::intron::maize psy::nos cassette was recovered with HindIII/PacI digestion and cloned into a binary vector containing a Gt::intron::SSUcrtI::nos cassette to give the construct 12421.

### 1.7 Maize psy (E1B) + crtI construct 12422

The vector with the Gt::intron::maize psy::nos cassette in the pPRP0117 backbone (from construction of 12421 above) was digested with the restriction enzymes flanking the intron followed by religation of the vector in order to remove the catalase intron. The Gt::maize psy::nos cassette was recovered with HindIII/PacI digestion and cloned into a binary vector containing a Gt::SSUcrtI::nos cassette to give the construct 12422.

### 1.8 Maize psy + crtI construct 12423

The Y1 maize *psy* (U32636) cds sequence was synthesised with NcoI and SfiI restriction sites added at the 5' and 3' end respectively. This was cloned into the NcoI and SfiI sites of a pPRP0117 based vector. The Gt::intron::maize psy::nos cassette was recovered

- 34 -

with HindIII/PacI digestion and cloned into a binary vector containing a Gt::intron::SSUcrtI::nos cassette to give the construct 12423.

### 1.9 Maize psy + crtI construct 12424

5 The vector with the Gt::intron::maize psy::nos cassette in the pPRP0117 backbone (from construction of 12423 above) was digested with the restriction enzymes flanking the intron followed by religation of the vector in order to remove the catalase intron. The Gt::maize psy::nos cassette was recovered with HindIII/PacI digestion and cloned into a binary vector containing a Gt::SSUcrtI::nos cassette to give the construct 12424.

10

### 2.0 Construction of vectors for plant transformation.

When providing vectors for plant transformation which utilise *Agrobacterium*, it is preferred that the sequences according to the invention are inserted between the border regions of a single TDNA region. *Agrobacterium* may be transformed in accordance with methods which are well known to the person skilled in the art, and/or via the methods disclosed herein.

15

### 3.0 Transformation of Rice

Based on the protocol published by Hiei et al (1994 The Plant Journal, 6 (2), 271-282).

20 The key modification involves use of a supervirulent strain in combination with a standard binary vector.

The basic procedure is as follows:

Mature rice seed (Asanohikari) are de-husked and surface sterilised by 70% ethanol for one minute followed by 4% Sodium hypochlorite + Tween for 30 minutes. Seed are then sown on a callus induction media (CIM) (N6 salts, N6 vitamins, 30 g/l sucrose, 1 g/l casein hydrolysate, 2 mg/l 2,4-D, pH 5.8, 4 g/l Gelrite) and placed in the dark at 30°C. After 3 weeks embryogenic calli are isolated and plated on CIM and placed under the same conditions. At the same time *Agrobacterium* cultures (AGL1 strain plus binary) are established by spreading inoculum onto LB plates plus Kanamycin (50 mg/l). After three days, *Agrobacterium* is scraped from the plate and re-suspended in AA1 + AS (AA salts, B5 vitamins, AA Amino Acids, 68.5 g/l sucrose, 36 g/l glucose, 0.5 g/l casein hydrolysate, 100 µM Acetosyringone, pH 5.2) to an optical density of 0.1 at 600 nm. The embryogenic calli is inoculated with the *Agrobacterium* solution for 10 minutes after

30

- 35 -

which the calli are spread on to plates of R2COMAS (R2 Micro salts, ½ R2 Macro salts, B5 Vitamins, 20 g/l sucrose, 10 g/l glucose, 1 g/l casein hydrolysate, 2 mg/l 2,4-D, 100 µM Acetosyringone, pH 5.2) and placed in the dark at 26°C. After three days calli are transferred on to selection media (N6 salts, N6 vitamins, 30 g/l sucrose, 1 g/l casein hydrolysate, 2 mg/l 2,4-D, 300 mg/l Timentin, 50 mg/l Hygromycin, 4 g/l Gelrite, pH 5.8) and place in the light at 30°C. All of the following steps occur under the same growth conditions (Light, 30°C). After three weeks the putative transgenic calli are transferred on a pre-regeneration media (N6 salts and vitamins, 30 g/l sucrose, 1 g/l casein hydrolysate, 1 mg/l 2,4-D, 300 mg/l Timentin, 50 mg/l Hygromycin, 6 g/l Gelrite, pH 5.8). After two weeks the good quality embryogenic calli is transferred to regeneration media (N6 Micro, ½ N6 Macro, N6 vitamins, AA amino acids, 20 g/l sucrose, 1 g/l casein hydrolysate, 0.2 mg/l NAA, 1 mg/l Kinetin, 50 mg/l Hygromycin, pH 5.8, Gelrite 6 g/l). After three weeks, plantlets that regenerate are subcultured to rooting media (½ MS salts, ½ B5 Vitamins, 10 g/l sucrose, 25 mg/l Hygromycin, pH 5.8, 8 g/l microagar. After two weeks the plantlets that have robust root systems are transferred to soil (50% John Innes #3, 50% peat, 26°C, 16 hour photoperiod) and covered with a fleece until the plants are established.

#### 4.0 Analysis of carotenoids in rice transformants

Carotenoids are extracted from seeds harvested at maturity. Seed is dehusked using a TR-200 Electromotion Rice Husker and then polished for 1 min. with a Pearlest polisher (Kett). Any white or discoloured seed are removed and 0.5g of the sample is ground for 2 minutes using a Glen Creston 8000 Mixer/Mill. A total of 1g of ground material/plant may be obtained by this method. This can be thoroughly mixed together prior to extraction of 2 x 0.5g portions of the powder. A standard compound can then be added to the samples for quantification of recoveries. Astaxanthin and echinenone are examples of standards which can be used. Samples are hydrated with 1ml of water and mixed using a vortex for a few seconds. 6ml of acetone is then added and the samples sonicated for 2 minutes. The samples are centrifuged for 5 min. at 3500 rpm. The supernatants are decanted, and the samples re-extracted twice more with 3 ml acetone, repeating the sonication/centrifugation steps between extractions. One extraction with 2 ml of tert-methylbutylether is then performed including the sonication/centrifugation steps and all supernatants for one sample pooled together. The total volume for each extract is



- 36 -

adjusted to 14 ml with acetone and the centrifugation step repeated. A 2ml aliquot of each sample is evaporated to dryness under a stream of nitrogen gas. These aliquots are re-dissolved in 75 µl of ethyl acetate, vortexed for 5-10 s and then transferred to amber HPLC insert vials. The vials are sealed immediately and centrifuged again at 3500 rpm prior to HPLC analysis.

The HPLC quantification is based on the response factor determined for each reference standards. The overall concentration of the prepared and dissolved reference standards is measured spectrophotometrically based on the published molar extinction coefficients and/or absorbance measured for solution of 1% concentration using 1cm path length (A1%1cm) (Ref: Britton G., Liaanen-Jensen S. and Pfander H. P. (1995) Carotenoids: Spectroscopy Vol 1B pp57-62, Birkhauser Verlag, Basel, ISBN 3-7643-2909-2). For each standard stock solution, the purity of the principal component is determined by HPLC. For components where no reference standard is available e.g. various cis isomers quantitative results are expressed using the response factor for that of β-carotene.

#### 4.1 HPLC Equipment & Conditions

Pump	:	Agilent 1100 Quaternary or Binary Pump; model number G1311A or G1312A, respectively
Degasser	:	Agilent 1100 Degasser; model number G1322A
Temperature Controlled Autosampler	:	Agilent 1100 Automatic Liquid Sampler; model number G1313A equipped with autosampler temperature controller model number G1330A
Detector	:	Agilent 1100 Diode array detector; model number G1315A or G1315B Agilent 1100 fluorescence detector; model number G1321A (for tocopherol analysis only).
Column Oven	:	Agilent 1100 Column Compartment; model number G1316A

- 37 -

Instrument Conditions	
Column	: YMC C30 3 $\mu$ m particles in 25 cm x 4.6 mm id stainless steel column + 1cm x 4 mm 5 $\mu$ m YMC C30 guard column
Column temperature	: 25°C
Sample temperature	: 4°C
Mobile phase	: Solvent A = MeOH/H <sub>2</sub> O/tert-butylmethylether (TBME) + 1.3mM NH <sub>4</sub> acetate (70/25/5 v/v) Solvent B = MeOH/H <sub>2</sub> O/TBME + 1.3mM NH <sub>4</sub> acetate (7/3/90 v/v)
Stop time	30 min
Post time	0 min
Flow rate	1 ml min <sup>-1</sup>
Injection volume	25 $\mu$ l in ethyl acetate

Gradient conditions (6%/min):

Time (min)	A % MeOH/H <sub>2</sub> O/TB ME + 1.3mM NH <sub>4</sub> acetate (70/25/5 v/v)	B % MeOH/H <sub>2</sub> O/TB ME + 1.3mM NH <sub>4</sub> acetate (7/3/90 v/v)
0	95	5
15.83	0	100
22	0	100
24	95	5
30	95	5

5

5.0 Results of HPLC quantification of carotenoids:

Results for rice plants transformed with construct 11586 (construct described in 1.5 above)

- 38 -

Sample identity	$\mu\text{g/g}$ dry weight (dwt) endosperm
wild type	0.05
11586-10	4.19
11586-7	6.11
11586-25	6.32
11586-15	7.55
11586-30	7.69
11586-28	9.05
11586-14	11.82
11586-20	12.82
11586-1	13.29
11586-12	18.59

Sequence ID Number	Components	$\mu\text{g/g}$ dwt endosperm
7	crtI + pepper psy	>3
8	crtI + tomato psy	>3
9	crtI + crtB	>5
-	Untransformed Control	0

#### 6.0 Transformation of rice – method used for transformation of rice with

5 *Agrobacterium* comprising constructs 12421, 12422, 12423 and 12424 (constructs described in examples 1.6, 1.7, 1.8 and 1.9 respectively).

For this example, rice (*Oryza sativa*) is used for generating transgenic plants. Various rice cultivars can be used (Hiei et al., 1994, Plant Journal 6:271-282; Dong et al., 1996, Molecular Breeding 2:267-276; Hiei et al., 1997, Plant Molecular Biology, 35:205-218).  
 10 Also, the various media constituents described below may be either varied in concentration or substituted. Embryogenic responses are initiated and/or cultures are established from mature embryos by culturing on MS-CIM medium (MS basal salts, 4.3 g/liter; B5 vitamins (200 x), 5 ml/liter; Sucrose, 30 g/liter; proline, 500 mg/liter;  
 15 glutamine, 500 mg/liter; casein hydrolysate, 300 mg/liter; 2,4-D (1 mg/ml), 2 ml/liter; adjust pH to 5.8 with 1 N KOH; Phytigel, 3 g/liter). Either mature embryos at the initial stages of culture response or established culture lines are inoculated and co-cultivated with the *Agrobacterium* strain LBA4404 containing the desired vector construction.

*Agrobacterium* is cultured from glycerol stocks on solid YPC medium (100 mg/L spectinomycin and any other appropriate antibiotic) for ~2 days at 28 °C. *Agrobacterium* is re-suspended in liquid MS-CIM medium. The *Agrobacterium* culture is diluted to an OD<sub>600</sub> of 0.2-0.3 and acetosyringone is added to a final concentration of 200 µM.

- 5 *Agrobacterium* is induced with acetosyringone before mixing the solution with the rice cultures. For inoculation, the cultures are immersed in the bacterial suspension. The liquid bacterial suspension is removed and the inoculated cultures are placed on co-cultivation medium and incubated at 22°C for two days. The cultures are then transferred to MS-CIM medium with Ticarcillin (400 mg/liter) to inhibit the growth of
- 10 *Agrobacterium*. For constructs utilizing the PMI selectable marker gene (Reed et al., In Vitro Cell. Dev. Biol.-Plant 37:127-132), cultures are transferred to selection medium containing Mannose as a carbohydrate source (MS with 2%Mannose, 300 mg/liter Ticarcillin) after 7 days, and cultured for 3-4 weeks in the dark. Resistant colonies are then transferred to regeneration induction medium (MS with no 2,4-D, 0.5 mg/liter IAA,
- 15 1 mg/liter zeatin, 200 mg/liter timentin 2% Mannose and 3% Sorbitol) and grown in the dark for 14 days. Proliferating colonies are then transferred to another round of regeneration induction media and moved to the light growth room. Regenerated shoots are transferred to GA7-1 medium (MS with no hormones and 2% Sorbitol) for 2 weeks and then moved to the greenhouse when they are large enough and have adequate roots.
- 20 Plants are transplanted to soil in the greenhouse and grown to maturity.

7.0 Method of extraction/quantification of carotenoids - for plants transformed in accordance with Example 6.0

- (a) Sample is grinded in a Geno/Grinder at 1600 rpm for 40 seconds.
- 25 (b) Representative amounts of homogenised sample (0.1 g) are weighed into an appropriate extraction vessel (e.g. 2 ml microcentrifuge tube). Sample weight is recorded.
- (c) Samples are hydrated with water (200 µl) and vortexed for about 2-3 seconds then left to stand for about 10 minutes.
- (d) Acetone (1.2 ml) is added to the sample.
- 30 (e) Ultrasonicate sample for 5 minutes.
- (f) Centrifuge sample for 3 min at 6000 rpm, then transfer the supernatant to another appropriately sized tube.
- (g) Steps (d) to (f) are repeated and combined with the previous extract.

- 40 -

- (h) Steps (d) to (f) are repeated with tert-butylmethylether (400  $\mu$ l) and combine with the acetone extracts.
- (i) All of the combined extracts are transferred into an appropriate size tube and evaporate it to dryness under a stream of nitrogen gas.
- 5 (j) Samples are redissolved in 2 ml of ethyl acetate + 0.5% BHT, vortex or ultrasonicate for 5-10 seconds and then centrifuged at 3500 rpm for 5 minutes before HPLC analysis.
- (k) Absorption at wavelength 450 nm is measured on a spectrophotometer for all samples and total carotenoid concentration is calculated based on an  $\epsilon$  value of 124865.

10 **8.0 Results of the method of Example 7.0**

Construct	Sample ID	Calculated Carotenoids ( $\mu$ g/g)
12421	RIGQ2003001046A4A	48.7
12421	RIGQ2003001063A43A	42.7
12421	RIGQ2003001063A4A	36.9
12421	RIGQ2003001063A99A	31.4
12421	RIGQ2003001063A59A	30
12421	RIGQ2003001063A75A	28.7
12421	RIGQ2003001048A3A	28.5
12421	RIGQ2003001050A23A	28
12421	RIGQ2003001050A13A	26.8
12421	RIGQ2003001048A25A	26.8
12421	RIGQ2003001048A61A	26.1
12421	RIGQ2003001049A46A	23.4
12421	RIGQ2003000993A63A	21.7
12421	RIGQ2003001049A43A	20.6
12421	RIGQ2003001049A26A	10.5
12421	RIGQ2003001050A18A	5.7

Construct	Sample ID	Calculated Carotenoids ( $\mu$ g/g)
12422	RIGQ2003001045A30A	58.8
12422	RIGQ2003001045A51A	54
12422	RIGQ2003000995A29A	43.6
12422	RIGQ2003000995A31A	43.1
12422	RIGQ2003001043A10A	42.4
12422	RIGQ2003001043A8A	42
12422	RIGQ2003000995A19A	41.9
12422	RIGQ2003000995A17A	40.9

- 41 -

Construct	Sample ID	Calculated Carotenoids ( $\mu\text{g/g}$ )
12422	RIGQ2003001060A23A	37.2
12422	RIGQ2003001052A56A	35
12422	RIGQ2003001051A75A	32
12422	RIGQ2003001045A68A	30.6
12422	RIGQ2003001045A86A	29.8
12422	RIGQ2003001045A49A	29.2
12422	RIGQ2003001060A25A	28.9
12422	RIGQ2003001051A64A	26.4
12422	RIGQ2003001051A34A	25.2
12422	RIGQ2003000994A7A	22.8
12422	RIGQ2003001045A74A	20.3
12422	RIGQ2003001051A46A	16.6
12422	RIGQ2003000995A7A	13.8
12422	RIGQ2003000995A26A	8.1
12422	RIGQ2003000995A41A	6.8
12422	RIGQ2003000995A8A	3.8

Construct	Sample ID	Calculated Carotenoids ( $\mu\text{g/g}$ )
12423	RIGQ2003001097A42A	52.80
12423	RIGQ2003001097A15A	50.40
12423	RIGQ2003001097A41A	48.00
12423	RIGQ2003001114A10A	44.20
12423	RIGQ2003001099A8A	40.00
12423	RIGQ2003001098A2A	28.80
12423	RIGQ2003001099A42A	26.00
12423	RIGQ2003001097A30A	24.00
12423	RIGQ2003001186A80A	23.30
12423	RIGQ2003001097A18A	21.60
12423	RIGQ2003001099A60A	17.60
12423	RIGQ2003001098A5A	17.40
12423	RIGQ2003001097A44A	12.00

Construct	Sample ID	Calculated Carotenoids ( $\mu\text{g/g}$ )
12424	RIGQ2003001121A60A	51.9
12424	RIGQ2003001121A20A	51.8
12424	RIGQ2003001093A11A	40
12424	RIGQ2003001094A85A	37.8
12424	RIGQ2003001093A58A	27.3

Construct	Sample ID	Calculated Carotenoids (µg/g)
12424	RIGQ2003001118A84A	26

Although the invention has been described by way of the above-referenced examples and the Sequence Listing and Figures as provided herein, it will be apparent that modifications and changes may be practiced which remain within the ambit of the present invention.

## References:

Name	Database Accession No.	Description	Reference
Glutelin1	NCBI D00584, EMBL Y00867	Rice, promoter of a rice seed storage protein	Takaiwa,F.; Ebinuma,H.; Kikuchi,S.; Oono,K.; Nucleotide sequence of a rice glutelin gene, FEBS Lett. 221:43 (1987)
Prolamin	D73383 D73384	Rice, promoter of a rice seed storage protein	Nakase,M.; Yamada,T.; Kira,T.; Yamaguchi,J.; Aoki,N.; Nakamura,R.; Matsuda,T.; Adachi,T.. The same nuclear proteins bind to the 5'-flanking regions of genes for the rice seed storage protein: 16kDa albumin, 13 kDa prolamin and type II glutelin. Plant Mol. Biol. 32:621-630 (1996)
Prolamin	M23746	Rice, promoter of a rice seed storage protein	Kim,W.T.; Okita,T.W.; Structure, expression, and heterogeneity of the rice seed prolamines Plant Physiol. 88:649 (1988)
1st intron of catalase gene	D21161	Castor bean	Suzuki,M.; Ario,T.; Hattori,T.; Nakamura,K.; Isolation and characterization of two tightly linked catalase genes from castor bean that are differentially regulated Plant Mol. Biol. 25:507 (1994)
SSU	X00806	Pea, signal peptide of small subunit of rubisco	Coruzzi,G.; Broglie,R.; Edwards,C.; Chua,N.H.; Tissue-specific and light- regulated expression of a pea nuclear gene encoding the small subunit of ribulose-1,5- bisphosphate carboxylase EMBO J. 3:1671 (1984)
SSU	X04334	Pea, signal peptide of small subunit of ribulose bisphosphate	Fluhr,R.; Moses,P.; Morelli,G.; Coruzzi,G.; Chua,N.H.; Expression dynamics of the pea rbcS multigene family and organ distribution of the transcripts EMBO J. 5:2063 (1986)

Name	Database Accession No.	Description	Reference
		carboxylase	
Maize psy	U32636	Maize phytoene synthase gene	Buckner,B.; Miguel,P.S.; Janick-Buckner,D.; Bennetzen,J.L.; The y1 gene of maize codes for phytoene synthase Genetics 143(1):479 (1996)
Pepper psy	X68017	Pepper phytoene synthase gene	Romer,S.; Hugueney,P.; Bouvier,F.; Camara,B.; Kuntz,M.; Expression of the genes encoding the early carotenoid biosynthetic enzymes in Capsicum annuum Biochem. Biophys. Res. Commun. 196:1414 (1993)
Tomato psy	Y00521	tomato phytoene synthase gene	Ray,J.; Bird,C.R.; Maunders,M.; Grierson,D.; Schuch,W.; Sequence of ptom 5 a ripening related cDNA from tomato Nucleic Acids Res. 15:10587 (1987)
Daffodil psy	X78814	daffodil phytoene synthase gene	Schledz,M.; Ali Babili,S.; Lintig,J.; Haubruck,H.; Rabbani,S.; Kleing,H.; Beyer,P.; Phytoene synthase from Narcissus pseudonarcissus: functional, expression, galactolipid requirement, topological distribution in chromoplasts and induction during flowering Plant J. 10:781 (1996)
CrtB	D90087	Erwinia uredovora phytoene synthase	Misawa,N.; Nakagawa,M.; Kobayashi,K.; Yamano,S.; Izawa,Y.; Nakamura,K.; Harashima,K.; Elucidation of the Erwinia uredovora carotenoid biosynthetic pathway by functional analysis of gene products expressed in Escherichia coli J. Bacteriol. 172:6704 (1990)
CrtI	D90087	Erwinia uredovora phytoene desaturase	Misawa,N.; Nakagawa,M.; Kobayashi,K.; Yamano,S.; Izawa,Y.; Nakamura,K.; Harashima,K.; Elucidation of the Erwinia uredovora carotenoid biosynthetic pathway by functional analysis of gene products expressed in Escherichia coli J. Bacteriol. 172:6704 (1990)
Nos	AJ237588	Agrobacterium tumefaciens Ti plasmid	
Phytoene Desaturase	M88683	Tomato	Giuliano,G.; Bartley,G.E.; Scolnik,P.A.; Regulation of carotenoid biosynthesis during tomato development Plant Cell 5(4):379 (1993)
Phytoene Desaturase	X68058	Pepper	Hugueney,P.; Roemer,S.; Kuntz,M.; Camara,B.; Characterization and molecular cloning of a flavoprotein catalyzing the



Name	Database Accession No.	Description	Reference
			synthesis of phytofluene and zeta-carotene in Capsicum chromoplasts Eur. J. Biochem. 209:399 (1992)
Phytoene Desaturase	U37285	Maize	Li,Z.H.; Matthews,P.D.; Burr,B.; Wurtzel,E.T.; Cloning and characterization of a maize cDNA encoding phytoene desaturase, an enzyme of the carotenoid biosynthetic pathway Plant Mol. Biol. 30(2):269 (1996)
Phytoene Desaturase	AF049356	Rice	Vigneswaran,A.; Wurtzel,E.T.; A Rice cDNA Encoding Phytoene Desaturase (Accession No. AF049356) (PGR99-131) Plant Physiol. 121(1):312 (1999)
ZDS	AF195507	Tomato	Bartley,G.E.; Ishida,B.K.; Zeta-carotene desaturase (Accession No. AF195507) from tomato (PGR99-181) Plant Physiol. 121(4):1383 (1999)
ZDS	X89897	Pepper	Albrecht,M.; Klein,A.; Hugueney,P.; Sandmann,G.; Kuntz,M.; Molecular cloning and functional expression in E. coli of a novel plant enzyme mediating zeta-carotene desaturation FEBS Lett. 372:199 (1995)
ZDS	AF047490	Maize	Luo,R.; Wurtzel,E.T.; A Maize cDNA Encoding Zeta Carotene Desaturase (Accession No. AF047490). (PGR99-118) Plant Physiol. 120(4):1206 (1999)
ZDS	AF054629	Rice	

## CLAIMS

1. A polynucleotide comprising:
  - (a) a region which comprises as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a nucleotide sequence derived from a bacterium which sequence encodes a carotene desaturase; and (iii) a transcription termination region; and
  - (b) a further region which comprises as operably linked components (i) a promoter which provides for seed preferred expression; and (ii) a nucleotide sequence encoding a phytoene synthase which sequence is derived from maize (*Zea sp.*) or rice (*Orzya sp.*); and (iii) a transcription termination region.
2. A polynucleotide according to claim 1 wherein the sequence which encodes the carotene desaturase is derived from *Erwinia sp.*
3. A polynucleotide according to claim 1 or claim 2 wherein said promoter is selected from the Glutelin 1 promoter and the Prolamin promoter and said transcription termination region is selected from the Nos; CaMV 35S and PotP1-II transcription termination regions.
4. A polynucleotide according to any one of claims 1 to 3 wherein the sequence which encodes carotene desaturase and the sequence which encodes phytoene synthase further comprises a sequence encoding a plastid targeting sequence.
5. A polynucleotide according to any one of claims 1 to 4 wherein said region and/or said further region further comprises an intron.
6. A polynucleotide according to any one of claims 1 to 5 which comprises a sequence selected from the group depicted as SEQ ID NO: 1; 2; 3; 4; and 6.
7. A polynucleotide sequence which is the complement of one which hybridises to a polynucleotide according to claim 6 at a temperature of about 65°C in a solution containing 6 x SSC, 0.01% SDS and 0.25% skimmed milk powder, followed by rinsing at the same temperature in a solution containing 0.2 x SSC and 0.1% SDS

wherein said polynucleotide sequence still comprises a region encoding a carotene desaturase and a further region encoding a phytoene synthase and when said polynucleotide sequence is inserted into plant material the seed of a plant regenerated from said material produce an increased amount of carotenoids when compared to a control like-seed.

8. A polynucleotide sequence according to claim 7 wherein when said polynucleotide sequence is inserted into plant material, the seed of a plant regenerated from said material produces at least a sixty fold increase in carotenoids when compared to a control like-seed.
9. A polynucleotide sequence according to claim 7 wherein when said polynucleotide sequence is inserted into plant material, the seed of a plant regenerated from said material produces at least a three hundred and fifty fold increase in carotenoids when compared to a control like-seed.
10. A polynucleotide sequence according to claim 7 wherein when said polynucleotide sequence is inserted into plant material the seed of a plant regenerated from said material produces carotenoids at a level of at least 10µg/g of endosperm of said seed.
11. A polynucleotide sequence according to claim 7 wherein when said polynucleotide sequence is inserted into plant material the seed of a plant regenerated from said material produces carotenoids at a level of at least 15µg/g of endosperm of said seed.
12. A polynucleotide sequence which is the complement of one which hybridises to a polynucleotide according to claim 6 at a temperature of about 65°C in a solution containing 6 x SSC, 0.01% SDS and 0.25% skimmed milk powder, followed by rinsing at the same temperature in a solution containing 0.2 x SSC and 0.1% SDS wherein said polynucleotide sequence still comprises a region encoding a carotene desaturase and a further region encoding a phytoene synthase and when said polynucleotide sequence is inserted into plant material the seed of a plant

regenerated from said material produce carotenoids amounting to at least 80% of the carotenoid content of a seed which comprises a polynucleotide selected from the group depicted as SEQ ID NO: 1; 2; 3; 4; 5 and 6.

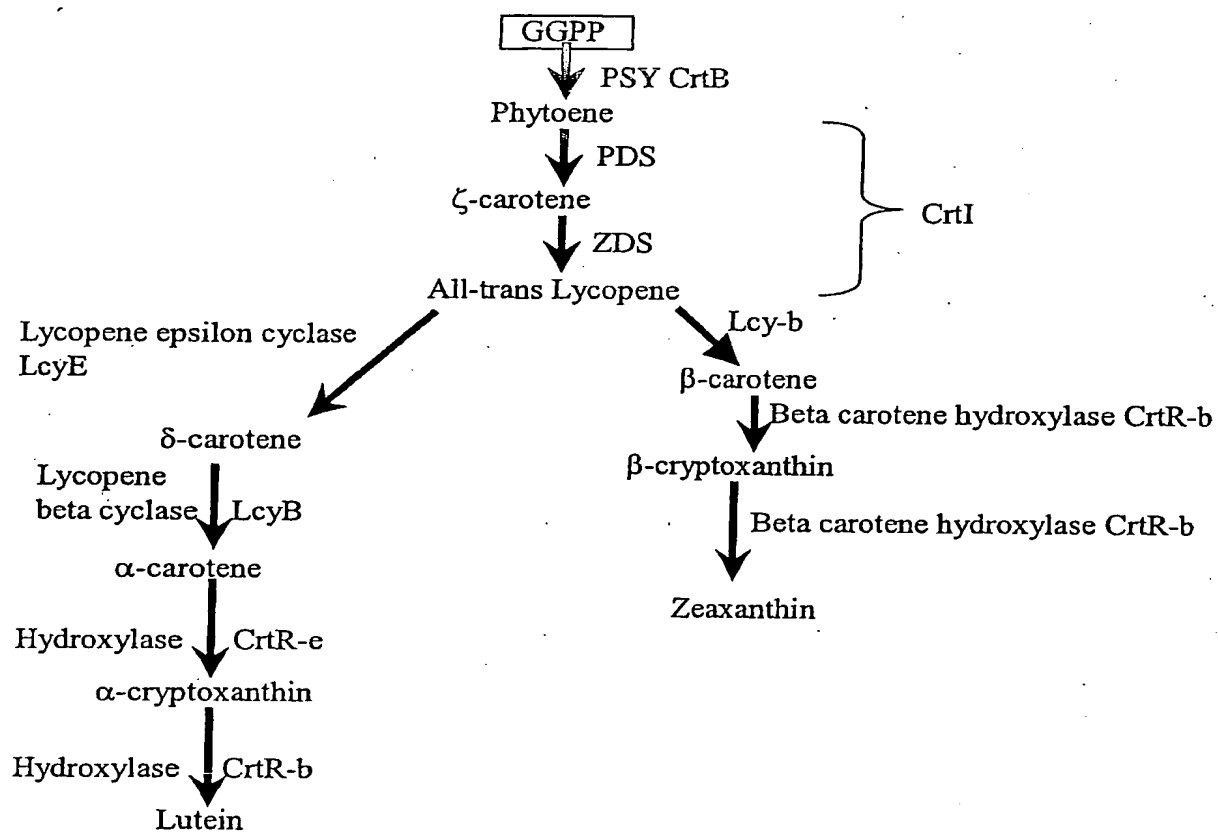
- 5 13. A polynucleotide sequence according to claim 12 wherein when said polynucleotide sequence is inserted into plant material the seed of a plant regenerated from said material produces carotenoids amounting to at least 100% of the carotenoid content of a seed which comprises a polynucleotide selected from the group depicted as SEQ ID NO: 1; 2; 3; 4; 5 and 6.
- 10 14. A polynucleotide sequence according to any one of claims 7 to 13 wherein said seed is a rice seed.
- 15 15. A polynucleotide or a polynucleotide sequence according to any one of claims 1 to 14 which further comprises a region which encodes a selectable marker.
16. A polynucleotide or a polynucleotide sequence according to claim 15 wherein said selectable marker comprises a mannose-6-phosphate isomerase gene.
- 20 17. A polynucleotide or a polynucleotide sequence according to any one of claims 1 to 16 which is codon optimised for expression in a particular plant species.
18. A polynucleotide or a polynucleotide sequence according to claim 17 wherein said plant species is rice (*Oryza sp.*).
- 25 19. A vector comprising a polynucleotide or a polynucleotide sequence according to any one of claims 1 to 18.
- 30 20. A method for increasing the carotenoid content of seeds comprising inserting into plant material a polynucleotide or a polynucleotide sequence according to any one of claims 1 to 18 or a vector according to claim 19; and regenerating a seed-containing plant from said material and identifying the seeds which contain carotenoids at levels greater than those of control like-seeds.

21. A method for increasing the carotenoid content of a seed comprising inserting into plant material a polynucleotide comprising a sequence selected from the group depicted as SEQ ID NO: 1; 2; 3; 4; 5 and 6 and regenerating a seed-  
5 containing plant from said material and identifying the seed which contains carotenoids at levels greater than those of a control like-seed.
22. A method according to claim 20 or claim 21 wherein said seed contains at least a sixty fold increase in carotenoids when compared to a control like-seed.  
10
23. A method according to claim 22 wherein said seed contains at least a three hundred and fifty fold increase in carotenoids when compared to control like-seed.
- 15 24. A method according to claim 20 or claim 21 wherein said seed contains carotenoids at a level of at least 10 $\mu$ g/g of endosperm of said seed.
25. A method according to claim 24 wherein said seed contains carotenoids at a level of at least 15 $\mu$ g/g of endosperm of said seed.  
20
26. A method according to any one of claims 20 to 25 wherein said carotenoids are selected from the group consisting of: lycopene; alpha-carotene; lutein; beta-carotene; zeaxanthin; antheraxanthin; violaxanthin; and neoxanthin or a combination thereof.
- 25
27. A seed obtained by a method according to any one of claims 20 to 26.
28. A seed according to claim 27 which is a rice seed.
- 30 29. A plant which comprises a seed according to claim 27 or claim 28.

30. A plant or plant material which comprises a polynucleotide or a polynucleotide sequence according to any one of claims 1 to 18 or a vector according to claim 19.
- 5 31. A plant or plant material according to claim 30 which is a rice plant or is rice plant material.
32. A plant or plant material according to claim 30 which is a maize plant or is maize plant material.
- 10 33. A plant according to any one of claims 29 to 32 which further comprises a polynucleotide which provides for a trait selected from the group consisting of: insect resistance and/or tolerance; nematode resistance and/or tolerance; herbicide resistance and/or tolerance; improved resistance and/or tolerance to stress; a substance having pharmaceutical activity; and any other desired agronomic trait.
- 15 34. Use of a polynucleotide or a polynucleotide sequence according to any one of claims 1 to 18 or a vector according to claim 19 in a method for the production of seeds containing increased carotenoids.
- 20 35. Use of a polynucleotide selected from the group depicted as SEQ ID NO: 1; 2; 3; 4; 5 and 6 for the production of a seed which contains carotenoids at levels greater than those of a control like-seed.
- 25 36. Use of a polynucleotide or a polynucleotide sequence according to any one of claims 1 to 18 or a vector according to claim 19 in a method for the production of a plant comprising said polynucleotide, said polynucleotide sequence or said vector.
- 30 37. Use of a polynucleotide selected from the group depicted as SEQ ID NO: 1; 2; 3; 4; 5 and 6 in a method for the production of a plant comprising said polynucleotide.

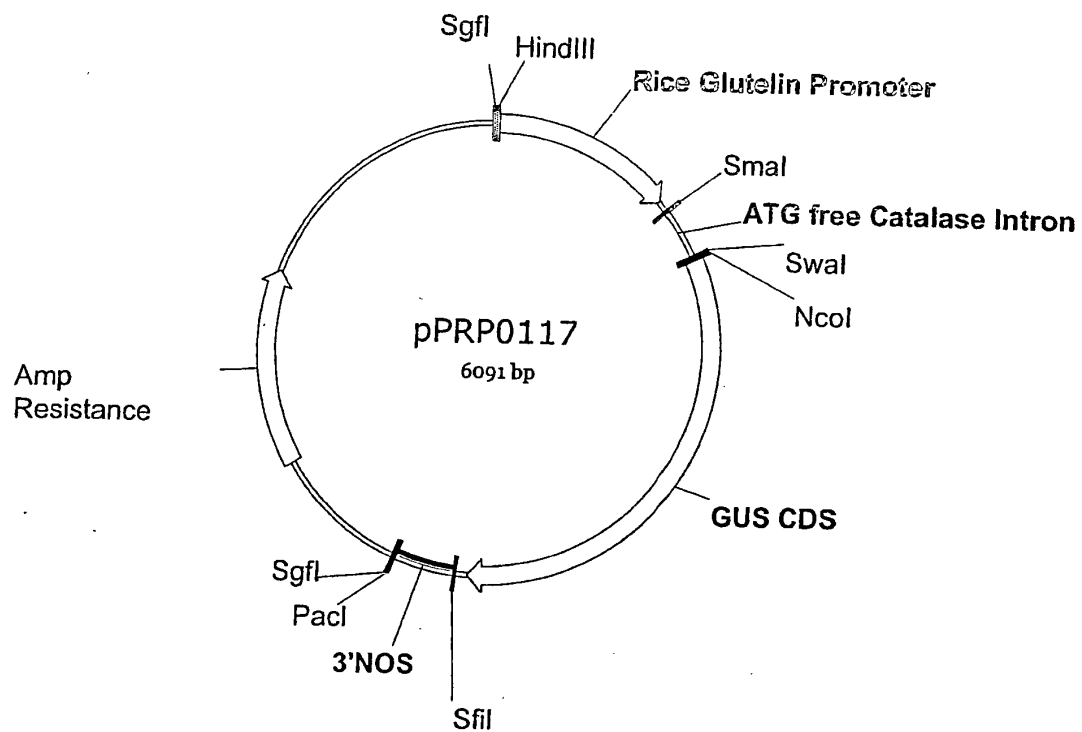
1/2

Figure 1



2/2

Figure 2





**10/549352**

-1-

**JC05 Rec'd PCT/PTO 14 SEP 2005**

## SEQUENCE LISTING

<110> SYNGENTA LIMITED

<120> ENHANCED ACCUMULATION OF CAROTENOIDS IN PLANTS

<130> 70237/WO

<150> US60/457,053

<151> 2003-03-24

<160> 38

<170> PatentIn version 3.1

<210> 1

<211> 5630

<212> DNA

<213> SYNTHETIC - 12423

<400> 1

gttaatcatg gtgtaggcaa cccaaataaa acaccaaagt atgcacaagg cagtttgttg	60
tattctgtag tacagacaaa actaaaagta atgaaagaag atgtggtgtt agaaaaggaa	120
acaatatcat gagtaatgtg tgagcattat gggaccacga aataaaaaga acattttgat	180
gagtcgtgta tcctcgatga gcctcaaaag ttctctcacc ccggataaga aacccttaag	240
caatgtgcaa agtttgcatt ctccactgac ataatgcaaa ataagatata atcgatgaca	300
tagcaactca tgcacatata catgcctctc tcaacctatt cattcctact catctacata	360
agtatcttca gctaaatgtt agaacataaa cccataagtc acgtttgatg agtattaggc	420
gtgacacatg acaaatcaca gactcaagca agataaagca aatgatgtg tacataaaac	480
tccagagcta tatgtcatat tgcaaaaaga ggagagctta taagacaagg catgactcac	540

-2-

aaaaattcat ttgcctttcg tgtcaaaaag      aggagggtt tacattatcc atgtcatatt  
600

gcaaaagaaa gagagaaaga acaacacaat gctgcgtcaa ttatacatat ctgtatgtcc      660  
atcattatcc atccaccttt cgtgtaccac acttcatata tcatgagtca cttcatgtct      720  
ggacattaac aaactctatc ttaacattta gatgcaagag cctttatctc actataaatg      780  
cacgatgatt tctcattgtt tctcacaaaa agcattcagt tcattagtcc tacaacaacg      840  
aattcggctt cccgggtaca gggtaaattt ctagtttttc tcttccattt tcttggttag      900  
gaccttttc tctttttatt tttttgagct ttgatcttcc tttaaactga tctatttttt      960  
aattgattgg ttatcgtgta aatattacat agctttaact gataatctga ttactttatt      1020  
tcgtgtgtct ttgatcatct tgatagttac agaaccgtcg actctagaga agccatttaa      1080  
atcgccgccca ccatggcttc tatgatatcc tcttccgctg tgacaacagt cagccgtgcc      1140  
tctagggggc aatccgccgc agtggtcca ttcggcgccc tcaaatccat gactggattc      1200  
ccagtgaaga aggtcaacac tgacattact tccattacaa gcaatggtgg aagagtaaag      1260  
tgcataaac caactacggt aattggtgca ggcttcggtg gcctggcact ggcaattcgt      1320  
ctacaagctg cggggatccc cgtcttactg cttgaacaac gtgataaacc cggcggtcgg      1380  
gcttatgtct acgaggatca ggggtttacc tttgatgcag gcccgacggt tatcaccgat      1440  
cccagtgccca ttgaagaact gtttgactg gcaggaaaac agttaaaaga gtatgtcgaa      1500  
ctgctgccgg ttacgccgtt ttaccgctg tgttgggagt cagggaaggt ctttaattac      1560  
gataacgata aaaccgggt cgaagcgcag attcagcagt ttaatccccg cgatgtcgaa      1620  
ggttatcgtc agtttctgga ctattcacgc gcggtgttta aagaaggcta tctgaagctc      1680  
ggtactgtcc cttttttatc gttcagagac atgcttcgcg ccgcacctca actggcgaaa      1740  
ctgcaggcat ggagaagcgt ttacagtaag gttgccagt acatcgaaga tgaacatctg      1800  
cgccaggcgt tttctttcca ctgctgttg gtggcgcca atcccttcgc cacctcatcc      1860  
atctatacgt tgatacacgc gctggagcgt gagtggggcg tctgggtttcc gcgtggcggc      1920  
accggcgcat tagttcagg gatgataaag ctgtttcagg atctgggtgg cgaagtcgtg      1980  
ttaaacgccca gattcagcca tatggaaacg acaggaaaca agattgaagc cgtgcattta      2040  
gaggacggtc gcaggttcct gacgcaagcc gtcgctcaa atgcagatgt ggttcatacc      2100  
tatecgacc tgttaagcca gcacctgcc gcggttaagc agtccaaca actgcagact      2160  
aagcgatga gtaactctct gtttgtgctc tattttggtt tgaatcacca tcatgatcag      2220  
ctcgcgcatc acacggtttg tttcgccccg cgttaccgcg agctgattga cgaaattttt      2280  
aatcatgatg gcctcgcaga ggactttctc ctttatctgc acgcgccctg tgtcacggat      2340

-3-

tcgtcactgg cgctgaagg ttgcggcagt tactatgtgt tggcgccggt gccgcattta 2400  
 ggcaccgcga acctcgactg gacgggttgag gggccaaaac tacgcgaccg tattttttgcg 2460  
 taccttgagc agcattacat gcctggctta cggagtcagc tggtcacgca ccggatgttt 2520  
 acgccgtttg attttcgcga ccagcttaat gcctatcatg gctcagcctt ttctgtggag 2580  
 cccgtttcta ccagagcgc ctggtttcgg ccgcataacc gcgataaaac cattactaat 2640  
 ctctacctgg tggcgcgagg cacgcacccc ggcgaggca ttcctggcgt catcggctcg 2700  
 gcaaaagcga cagcagggtt gatgctggag gatctgattt gaggccatgc aggccgatcc 2760  
 ccgatcgttc aaacatttgg caataaagtt tcttaagatt gaatcctgtt gccggtcttg 2820  
 cgatgattat catataattt ctgttgaatt acgttaagca tgtaataatt aacatgtaat 2880  
 gcatgacgtt atttatgaga tgggttttta tgattagagt cccgcaatta tacatttaat 2940  
 acgcgataga aaacaaaata tagcgcgcaa actaggataa attatcgcg cgggtgtcat 3000  
 ctatgttact agatcggggc ttaataagct tgttaatcat ggtgtaggca acccaaataa 3060  
 aacaccaaaa tatgcacaag gcagtttggt gtattctgta gtacagacaa aactaaaagt 3120  
 aatgaaagaa gatgtggtgt tagaaaagga aacaatatca tgagtaatgt gtgagcatta 3180  
 tgggaccacg aaataaaaag aacattttga tgagtcgtgt atcctcgatg agcctcaaaa 3240  
 gttctctcac cccggataag aaacccttaa gcaatgtgca aagtttgcat tctccactga 3300  
 cataatgcaa aataagatat catcgatgac atagcaactc atgcatcata tcatgcctct 3360  
 ctcaacctat tcatctctac tcatctacat aagtatcttc agctaaatgt tagaacataa 3420  
 acccataagt cacgtttgat gagtattagg cgtgacacat gacaaatcac agactcaagc 3480  
 aagataaagc aaaatgatgt gtacataaaa ctccagagct atatgtcata ttgcaaaaag 3540  
 aggagagctt ataagacaag gcatgactca caaaaattca tttgcctttc gtgtcaaaaa 3600  
 gaggagggct ttacattatc catgtcatat tgcaaaagaa agagagaaag aacaacacaa 3660  
 tgctgcgtca attatacata tctgtatgtc catcattatt catccacctt tcgtgtacca 3720  
 cacttcatat atcatgagtc acttcatgtc tggacattaa caaactctat cttaacattt 3780  
 agatgcaaga gcctttatct cactataaat gcacgatgat ttctcattgt ttctcacaaa 3840  
 aagcattcag ttcattagtc ctacaacaac gaattcggct tcccgggtac agggtaaatt 3900  
 tctagttttt ctcttcatt ttcttggtta ggacctttt ctctttttat ttttttgagc 3960  
 tttgatcttt ctttaaaactg atctattttt taattgattg gttatcgtgt aaatattaca 4020  
 tagctttaac tgataatctg attactttat ttcgtgtgtc tttgatcatc ttgatagtta 4080  
 cagaaccgtc gactctagag aagccattta aatcgccgcc accatggcca tcatactcgt 4140

-4-

acgagcagcg tcgccggggc tctccgccgc cgacagcadc agccaccagg ggactctcca 4200  
gtgctccacc ctgctcaaga cgaagaggcc ggcggcgcg cggtggatgc cctgctcgct 4260  
ccttggcctc cacccgctggg aggctggccg tccctcccc gccgtctact ccagcctgcc 4320  
cgtcaaccgc gcgggagagg ccgtcgtctc gtccgagcag aaggtctacg acgtcgtgct 4380  
caagcaggcc gcattgctca aacgccagct gcgcacgcc gtcctcgacg ccaggcccca 4440  
ggacatggac atgccacgca acgggctcaa ggaagcctac gaccgctgcg gcgagatctg 4500  
tgaggagtat gccaaagcgt ttacctcgg aactatgttg atgacagagg agcggcgccg 4560  
cgccatatgg gccatctatg tgtggtgtag gaggacagat gagctttag atgggcaaaa 4620  
cgccaactac attacaccaa cagctttgga ccggtgggag aagagacttg aggatctgtt 4680  
cacgggacgt ccttacgaca tgcttgatgc cgtctctct gataccatct caaggttccc 4740  
catagacatt cagccattca gggacatgat tgaagggatg aggagtgatc ttaggaagac 4800  
aaggataaac aacttcgacg agctctacat gtactgctac tatgttgctg gaactgtcgg 4860  
gttaatgagc gtacctgtga tgggcatcgc aaccgagtct aaagcaacaa ctgaaagcgt 4920  
atacagtgtc gccttggtc tgggaattgc gaaccaactc acgaacatac tccgggatgt 4980  
tggagaggat gctagaagag gaaggatata ttaccacaa gatgagcttg cacaggcagg 5040  
gctctctgat gaggacatct tcaaaggggt cgtcacgaac cgggtggagaa acttcatgaa 5100  
gaggcagatc aagagggcca ggatgttttt tgaggaggca gagagagggg taactgagct 5160  
ctcacaggct agcagatggc cagtatgggc ttccctgttg ttgtacaggc agatcctgga 5220  
tgagatcgaa gccaacgact acaacaactt cacgaagagg gcgtatgttg gtaaagggaa 5280  
gaagttgcta gcacttcctg tggcatatgg aaaatcgcta ctgctcccat gttcattgag 5340  
aaatggccag acctagggcc atgcaggccg atccccgac gttcaaacaat ttggcaataa 5400  
agtttcttaa gattgaatcc tgttgccggt cttgcgatga ttatcatata atttctgttg 5460  
aattacgtta agcatgtaat aattaacatg taatgcatga cgttatttat gagatgggtt 5520  
tttatgatta gagtcccgca attatacatt taatacgca tagaaaacaa aatatagcgc 5580  
gcaaactagg ataaattatc gcgcgcggtg tcatctatgt tactagatcg 5630

&lt;210&gt; 2

&lt;211&gt; 5630

&lt;212&gt; DNA

&lt;213&gt; SYNTHETIC - 12421